Molecular and biochemical characterization of subtilisin-like proteases in arabidopsis thaliana

Hamilton, John Michael Uwe

Molecular and biochemical characterization of subtilisin-like proteases in *Arabidopsis thaliana*

A thesis submitted to the

University of Durham

for the degree of Doctor of Philosophy

2000

John Michael Uwe Hamilton

Department of Biological Sciences

University of Durham

Supervisors: Prof A. R. Slabas
Dr. W. Schuch

Abstract

Subtilisin-like proteases form a large group of serine proteases with diverse functions, including the specific processing of a variety of proproteins and prohormones, and are found in prokaryotic and eukaryotic organisms. The work in this thesis focuses mainly on the Ara12 subtilisin-like protease following its discovery in the filtrate of *Arabidopsis* cell suspension cultures.

Evidence obtained by Southern blotting and database searching is presented for the existence of a large gene family encoding subtilisin-like proteases in the model plant *Arabidopsis thaliana*. There may be more than fifty members in this gene family. Three of the corresponding DNA sequences have been cloned by RT-PCR and used as probes in Northern analysis to investigate the tissue specificity of the gene transcripts. These three genes appear to be expressed to varying degrees in *Arabidopsis* leaf, stem, root and silique tissues.

A 650bp cDNA fragment encoding the C-terminal portion of the Ara12 protease has been obtained by RT-PCR, ligated to the *malE* gene and overexpressed as a fusion protein in *E. coli* cells. Polyclonal antisera have been raised against a combination of the fusion protein and the Ara12 C-terminal protein purified after cleavage from the fusion protein using Factor Xa protease. Ara12 protein has been detected in *Arabidopsis* tissues, particularly in siliques and stems, by Western blotting using these antibodies. An apoplastic location has been ascribed to Ara12 protease by immunocytochemistry using electron microscopy.

The mature Ara12 subtilisin-like protease has been purified to homogeneity from *Arabidopsis* cell suspension culture filtrate by ion exchange chromatography and hydrophobic interaction chromatography. The purified enzyme has an acidic pH optimum of approximately pH5.5, which is unusual for a plant subtilisin-like protease. Ara12 protease is relatively thermostable and is activated in the presence of Ca^{2+} ions. The known serine protease inhibitors phenylmethanesulphonyl fluoride (PMSF), 4-(2-aminoethyl) benzenesulphonyl fluoride (AEBSF) and diisopropyl fluorophosphate (DFP) have an inhibitory effect on the proteolytic activity of Ara12. Substrate specificity studies have been performed using artificial peptide substrates, native proteins and cell wall protein extracts from *Arabidopsis* cells.

Table of Contents

Chapter 4: Generation of polyclonal antibodies against and immunolocalisation of Ara12/Slpa subtilisin-like protease

Chapter 5: Purification of Ara12/Slpa subtilisin-like protease from *Arabidopsis* cell suspension culture filtrate

Chapter 6: Biochemical analysis of Ara12/Slpa protease

List of Figures

List of Tables

Declaration

No part of the work contained within this thesis has been previously submitted for any other degree or qualification at the University of Durham or any other university or institute of learning. All the work was performed by the author, except where stated otherwise.

Statement of Copyright

Acknowledgements

I would like to thank my supervisor at the University of Durham, Professor A. R. Slabas, for advice throughout this work and help in the laboratory. I am grateful to Dr. Wolfgang Schuch and thankfully acknowledge a CASE award from Zeneca Agrochemicals.

I am particularly grateful to Dr. Duncan Robertson and Dr. Annette Carey for providing general advice and assistance. I would like to thank Bill Simon and Dr. Helen Denton for help with protein purification and general discussions.

I would also like to thank Andrew Dinsmore and Jane Bird at AstraZeneca for arranging the inoculations of antigen during the generation of anti-Ara12 antisera. I would especially like to thank Stefan Hyman for his work on the EM immunolocalisation of Ara12 in *Arabidopsis* tissue, John Gilroy for the sequencing of the Ara12 protein, Vicky Kelly and Gillian Storey for nucleotide sequencing and David Hutchinson and Paul Sidney for photography. Thanks also to Dr. Dave Bown for help with protease assays and everyone who worked in and around my laboratory.

Finally, I would like to thank my family for always being there for me and supporting me.

Abbreviations

A_{280}	absorbance of light of wavelength 280nm
A	adenine
AEBSF	4 - (2-aminoethyl) benzenesulphonyl fluoride
ATP	adenosine 5'-triphosphate
AMC	7-amido-4-methylcoumarin
Boc	*t*-butyloxycarbonyl
bp	base pairs
BSA	bovine serum albumin
C	cytosine
cAMP	cyclic adenosine 3',5'-monophosphate
CAPS	3-[cyclohexylamino]-1-propanesulphonic acid
cDNA	complementary deoxyribonucleic acid
CDTA	cyclohexane diaminotetraacetic acid
Ci	Curie
2,4-D	2,4-dichlorophenoxyacetic acid
DEPC	diethylpyrocarbonate
DFP	diisopropyl fluorophosphate
dNTP	2'-deoxyribonucleoside 5'-triphosphate
ddNTP	2',3'-dideoxyribonucleoside 5'-triphosphate
DMF	dimethyl formamide
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ECL	enhanced chemiluminescence
EDTA	ethylenediamine tetraacetic acid
EGTA	[ethylenebis(oxyethylenenitrilo)] tetraacetic acid
EtBr	ethidium bromide
FITC	fluorescein isothiocyanate
g	acceleration due to gravity
G	guanine
Glp	pyroglutamyl
HEPES	*N*-[2-hydroxyethyl]piperazine-*N*'-[2-ethane-sulphonic acid]
HPLC	high performance liquid chromotography
IgG	immunoglobulin G
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobases

kDa	kilodaltons
K_M	Michaelis constant
LMP	low melting point
MES	2-[*N*-morpholino] ethanesulphonic acid
Mr	molecular weight
mRNA	messenger ribonucleic acid
N_n	nucleotide repeated n times
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming units
pI	isoelectric point
PMSF	phenylmethanesulphonyl fluoride
pNA	*p*-nitroanilide
psi	pounds per square inch
RNA	ribonucleic acid
rpm	revolutions per minute
RT-PCR	reverse transcription polymerase chain reaction
SDS	sodium dodecyl sulphate
T	thymine
TBST	Tris buffered saline, Tween
TCA	trichloroacetic acid
Tm	melting temperature
TEMED	*N, N, N', N'* tetramethylethylenediamine
Tris	2-amino,2-(hydroxymethyl)propane 1,3-diol
U	units of enzyme activity
UV	ultraviolet
V_{max}	maximum velocity
v/v	volume to volume
W	watts
w/v	weight to volume
X-gal	5-bromo,4-chloro,3-indoyl, β-D-galactopyranoside
Z	benzyloxycarbonyl

Chapter 1

Introduction

1.1 Proteases and the role of proteolysis

Proteases, proteinases or peptidases, are defined as a group of enzymes which are capable of hydrolysing peptide bonds. Most of these enzymes are involved in proteolysis, that is the enzymatic hydrolysis of proteins into more simple components (polypeptides, peptides and amino acid residues), however some members of this group of enzymes can cleave small peptides, but do not act on proteins. Many proteases can also cleave esters, amides and β-lactams, as well as peptide bonds. The terms protease, proteinase and peptidase are all now generally used to refer to this large group of enzymes involved in the hydrolysis of peptide bonds, although historically they have had different meanings (Barrett and McDonald, 1986). The term protease will mostly be used here, as it is extremely common in the comprehensive literature associated with these enzymes, although workers involved in the task of classification are increasingly recommending the use of the term peptidase, due to its more specific nature.

Proteases are found in almost all biological cells, tissues and fluids and are important as they carry out fundamental modifications of proteins. Proteolytic modification can range from limited protein cleavage, or processing, to total protein degradation. Other important modifications of proteins include phosphorylation, glycosylation, sulphation and α-amidation of amino acid residues.

Depending on their site of action, two main classes of protease can be distinguished: endopeptidases and exopeptidases (Barrett, 1994). Exopeptidases are proteolytic enzymes which act at the ends of peptide chains, whereas endopeptidases act preferentially at sites away from the peptide chain termini (see Table1.1). Endo-peptidases generally cleave bonds in the inner parts of peptide chains of proteins,

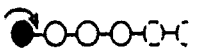
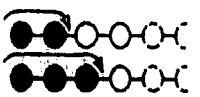
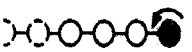
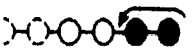
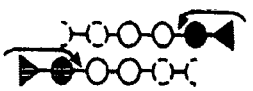
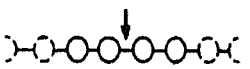
Action	Group	EC subsection
	Exopeptidases	
	Aminopeptidases	3.4.11
	Dipeptidyl-peptidases, tripeptidyl-peptidases	3.4.14
	Carboxypeptidases	3.4.16–18
	Peptidyl-dipeptidases	3.4.15
	Dipeptidases	3.4.13
	Omega peptidases	3.4.19
	**Endopeptidases**	3.4.21–24 and 99

Table 1.1 Classification of peptidases by type of reaction catalyzed. Open circles represent amino acid residues and filled circles are the residues comprising the blocks of one, two or three terminal amino acids that are cleaved off by these enzymes. The triangles indicate the blocked termini that provide substrates for some of the omega peptidases. The EC classification of the enzymes is also given. Figure taken from Barrett, 1994.

although oligopeptidases, which are also classed as endopeptidases, act specifically on shorter oligopeptides and polypeptides.

Exopeptidases can act at either end of peptide chains. Exopeptidases which operate at the N-terminus of peptide chains can release a single amino acid residue, a dipeptide or a tripeptide, and are known as aminopeptidases, dipeptidyl-peptidases and tripeptidyl-peptidases, respectively. Similarly, exopeptidases can act at the C-terminus of peptide chains and release a single amino acid residue or a dipeptide, and these are known, repectively, as carboxypeptidases and peptidyl-dipeptidases. Dipeptidases are exopeptidases which specifically cleave dipeptides only, whilst omega peptidases are exopeptidases which remove terminal amino acid residues which have been substituted, cyclized or linked by isopeptide bonds (peptide bonds other than α-carboxyl to α-amino group bonds).

The essential role of proteases in the regulation of cellular processes is increasingly becoming apparent. One of the most important of these is cellular housekeeping. Abnormal proteins which arise as a result of mutations or under conditions of stress or disease are degraded in proteolytic pathways (Maurizi, 1992). Prohormones, receptors, antigens and many other effectors are processed by proteases. Proteases also form an integral part of the protein targeting machinery, where they act as signal peptidases. Many enzymes are regulated by proteolysis, as degradation can offer a precise and rapid way of controlling key enzymatic steps. Examples of this include numerous metabolic enzymes, transcription factors, protein kinases and cell division cycle regulators, such as cyclins (Glotzer *et al.*, 1991). Proteins are broken down to meet the nutritional requirements of the cell. During seed germination in plants, proteins deposited in protein bodies are degraded to supply growing cells with amino acids. In animals, proteins are broken down by proteases in lysosomes in response to starvation

(Dice, 1987). Protein degradation is thought to occur in vacuoles in plants and fungi (Vierstra, 1993, 1996). Proteolytic enzymes are also involved in another important process known as programmed cell death (PCD), which is a selective method of destroying unwanted cells. This might occur because the cells or tissues are undergoing normal developmental changes and are no longer required or because the tissue becomes specialized. In plants this occurs in the differentiation of the xylem and sclerenchyma, in tapetal cells during microsporogenesis, in aleurone cells, in root cap cells and during leaf and flower senescence (see Pennell and Lamb, 1997 for review). Cells which have been subjected to biotic or abiotic stress, especially damaged cells, can also undergo PCD. Cell death occurs for instance in plants which are under attack from avirulent pathogens and react with a localized hypersensitive response (HR) (Lamb and Dixon, 1994). There is evidence that the hypersensitive response is a form of PCD (Pennell and Lamb, 1997; D'Silva *et al.*, 1998).

Endopeptidases are classified according to features of their principal catalytic sites into aspartic (or aspartyl) proteases, cysteine proteases (known previously as thiol proteases), serine proteases and metallo-proteases. A further catalytic type of endopeptidase has only been recognised relatively recently: the threonine peptidases (Brannigan *et al.*, 1995). The nucleophile in the catalytic site of the serine, threonine and cysteine proteases consists of part of an amino acid residue, whereas it is an activated water molecule in the aspartic and metallo-proteases. A large number of proteases have not yet been classifed, due to a lack of information regarding their primary and tertiary structures.

1.2 Serine proteases

The first serine proteases to be extensively studied were the pancreatic digestive enzymes chymotrypsin, trypsin and elastase (Shotton, 1971). Once activated, these enzymes are involved in the degradation of dietary protein into peptides and amino acid residues. Much of the current knowledge about protein structure and function, as well as protease catalytic mechanisms, was obtained as a result of examining these and other serine proteases, such as subtilisin (Wells *et al.*, 1987).

Serine proteases are involved in diverse biological processes, such as blood clotting and dissolution (Neurath, 1986), the processing of protein hormones, receptors and growth factors (Nakayama, 1997) and fertilization (Stambaugh *et al.*, 1969), as well as digestive processes.

The serine proteases have been classified into the following six superfamilies, or clans, depending on their tertiary structures and the order of their catalytic residues: (chymo)trypsin-like peptidases (clan SA), subtilisin-like peptidases (clan SB), peptidases containing the α/β hydrolase fold (clan SC), serine-type *D*-Ala-*D*-Ala peptidases (clan SE), peptidases containing the Ser/Lys catalytic dyad (clan SF) and the herpes virus assemblins (clan SH) (Rawlings and Barrett, 1994). Proteins containing the characteristic α/β hydrolase fold have a tertiary structure consisting of parallel β sheets comprised of β/α/β units. A number of serine protease families are as yet unassigned, as the order of their catalytic residues and their tertiary structures are not known. These have temporarily been assigned to clan SX until they can be properly classified (Rawlings, 1997).

A clan, or group of families, is thought to originate from a common single ancestral protein. The (chymo)trypsin-like superfamily and the subtilisin-like (or subtilase)

superfamily are the two largest of these clans (Siezen and Leunissen, 1997). Enzymes from both the chymotrypsin-like and subtilisin-like clans have strikingly similar active sites, but judging by their tertiary structures and the amino acid residues surrounding the catalytic sites, it appears that members of the two different clans arose via different evolutionary routes (Polgár, 1987). This is a prime example of convergent evolution, as it suggests that these serine proteases have evolved in parallel arriving at a similar way to provide organisms with a selective advantage. The active sites of chymotrypsin-like and subtilisin-like enzymes contain a so-called catalytic triad consisting of an aspartate, a histidine and a serine residue. A common catalytic mechanism was discovered for these enzymes (Blow *et al.*, 1969), whereby the active site serine residue is hydrogen bonded to the imidazole side chain of the active site histidine residue. The latter is also hydrogen bonded to the active site apartate residue, completing a charge relay system (see Figure 1.1). This allows the charge of the buried aspartate residue to be conveyed within this system, causing the normally inert serine hydroxyl group to become highly nucleophilic. In this way the active site serine hydroxyl group is able to initiate a nucleophilic attack on the carbonyl carbon atom of the peptide bond to be cleaved.

1.3 Subtilisins

The enzyme subtilisin was first isolated from the bacterium *Bacillus subtilis* (Ottensen and Svenson, 1970; Markland and Smith, 1971), from which it derived its name, although the first enzyme of this type, subtilisin Carlsberg, was discovered in *Bacillus licheniformis* (Linderstrøm-Lang and Ottensen, 1947). The term subtilisin is

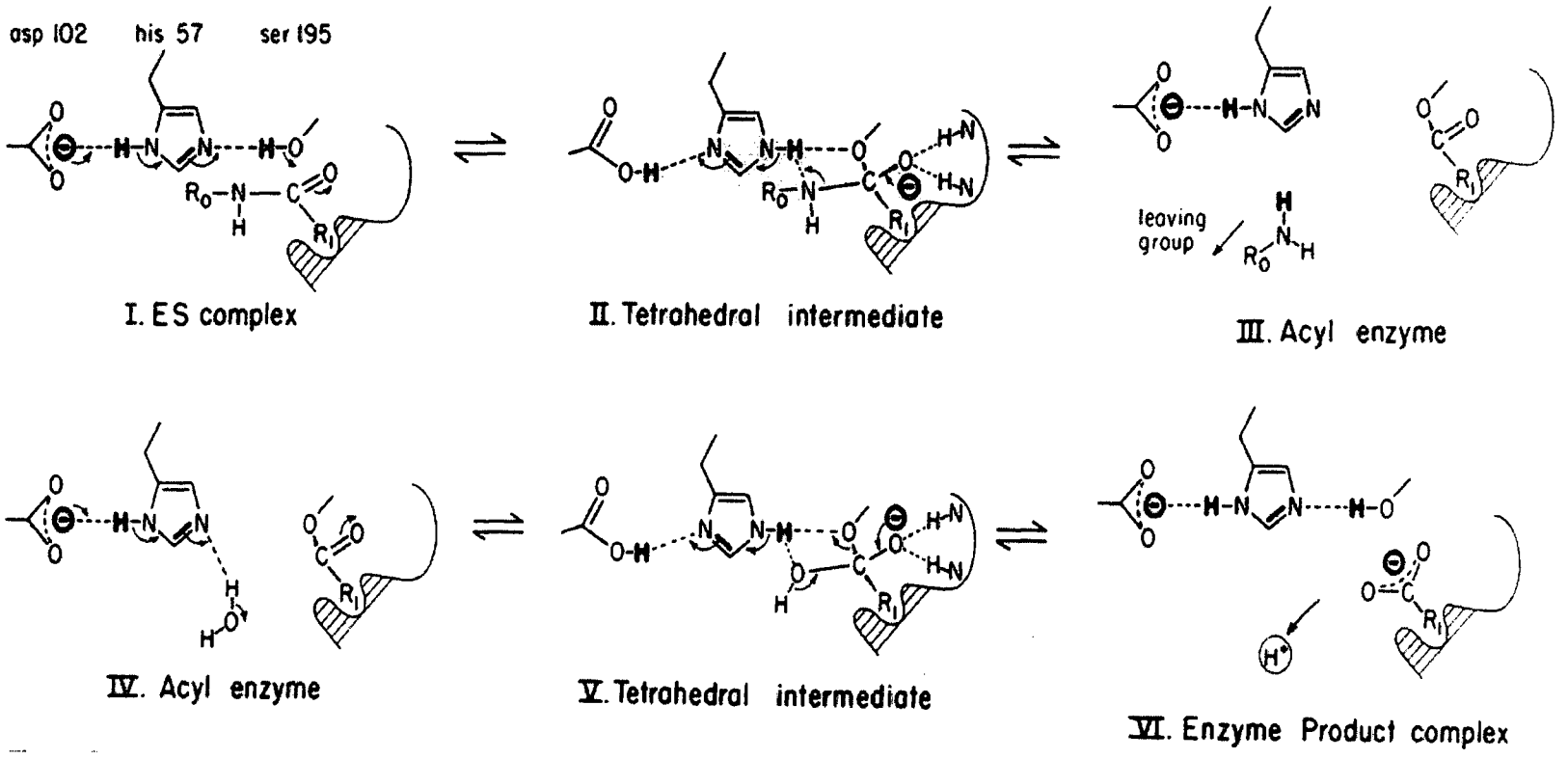


Figure 1.1 The mechanism of serine protease hydrolysis of peptide or amide bonds as exemplified by chymotrypsin. In this representation the catalytic triad, consisting of an aspartate residue (asp 102), a histidine residue (his 57) and a serine residue (ser 195), is shown. Figure taken from Stroud *et al.*, 1975.

now used for homologous enzymes isolated from any *Bacillus* species. Various subtilisins have been isolated and investigated. The subtilisins are extracellular serine proteases which have been heavily studied on account of their use as additives to detergents. This has made them the most commercially valuable group of enzymes. Although subtilisins are mainly used in washing powders, they have also been used in a variety of different processes from cleaning contact lenses to industrial scale biocatalysis.

Subtilisins have been found to possess a broad substrate specificity and are stable in alkaline conditions (Betzel *et al.*, 1992). They are monomeric enzymes and have two Ca^{2+} binding sites involved in stabilizing their structure (Pantoliano *et al.*, 1988). Various properties of the subtilisins have been altered by protein engineering (Wells and Estell, 1988). These include altering their stability to different pH (Russell and Fersht, 1987), temperature (Pantoliano *et al.*, 1989) and non-aqueous conditions (Chen and Arnold, 1993), as well as altering their substrate specificity (Rheinnecker *et al.*, 1994; Perona and Craik, 1995; Ballinger *et al.*, 1996), disulphide bond content (Wells and Powers, 1986) and their ability to resist oxidative inactivation (Estell *et al.*, 1985). Incorporating cysteine residues into subtilisin protein at suitable sites to form disulphide bridges, successfully increased the heat stability of the enzyme, making it more resistant to higher operating temperatures (Pantoliano *et al.*, 1987).

Another interesting finding has been that if the Ser221 residue of subtilisin BPN' (subtilisin from *Bacillus amyloliquefaciens*) is mutated to a cysteine residue and the Pro225 residue is mutated to an alanine residue, then the double mutant has very unusual properties. The mutant enzyme (called subtiligase) can ligate peptide derivatives site-specifically to the N-termini of peptides or proteins (Chang *et al.*, 1994). In this way it has been possible to recreate synthetic ribonuclease A variants

from blocks of artificial peptides (Jackson *et al.*, 1994). The Ser221 residue can also be chemically modified using PMSF and hydrogen selenide resulting in a seleno-cysteine mutant enzyme. This seleno-subtilisin also acts as a peptide ligase under anaerobic conditions (Wu and Hilvert, 1989).

Although the biochemistry of these proteases has been well-studied, there is some confusion about their biological role. Subtilisins and neutral metallo-proteases are secreted into the extracellular matrix of bacilli. Initially these enzymes were believed to function in the formation of heat-resistant endospores, a process which is usually initiated in these bacteria under conditions of nutrient deprivation (Piggot and Coote, 1976). This view was held because the activities of these (and many other enzymes) increased during sporulation, and a temperature-sensitive mutant of *B. subtilis,* which produced an inactive subtilisin at a certain temperature, was found to be asporogenous at this restrictive temperature (Leighton *et al.*, 1973). However it remained unclear whether or not the activity of the serine protease was directly connected with sporulation (Priest, 1977).

Using different *Bacillus subtilis* strains, an inverse relationship between the extracellular proteolytic activity and the turnover rate of cell wall peptidoglycan has been demonstrated (Jolliffe *et al.* 1980). Addition of subtilisin to a protease-deficient strain decreased the rate of cell wall turnover. Conversely, addition of the serine protease inhibitor PMSF to hyperprotease-producing strains substantially increased their respective rates of cell wall turnover. These results suggested that cell wall turnover could be regulated by subtilisin.

However, subsequent studies using *in vitro*-derived deletion mutants of subtilisin and the neutral metallo-protease indicated that these enzymes have no effect on sporulation, morphology or growth (Stahl and Ferrari, 1984; Yang *et al.*, 1984).

Therefore the extracellular proteases might only have a role in nutrition as scavenger enzymes of extracellular proteins, by hydrolysing them to peptides and amino acids, although no direct proof for this has been offered.

Although first examined in *Bacillus* bacteria, proteins homologous to the subtilisins have been found in all kingdoms: archaea, bacteria, fungi, and in higher eukaryotes including plants, insects, nematodes, molluscs, amphibians, fish and mammals. These subtilisin-like serine proteases have collectively been termed subtilases (Siezen *et al.*, 1991). The amino sequences of the catalytic domains of these enzymes were compared by multiple sequence alignment. Analysis of these data has led to a classification of the subtilases into the following six families: subtilisin family, thermitase family, proteinase K family, lantibiotic peptidase family, kexin family and pyrolysin family (Siezen and Leunissen, 1997). Figure 1.2 shows this classification of the subtilase (or subtilisin-like protease) superfamily based on catalytic domain sequence comparisons.

Proteases from the subtilisin and thermitase families have only been found in micro-organisms, whereas enzymes from the proteinase K family occur in micro-organisms and in fungi. Lantibiotic peptidases are found exclusively in Gram positive bacteria, and are responsible for removing the leader peptides from the precursors of the antimicrobial peptides known as lantibiotics.

The kexin family is comprised of a large group of subtilisin-like proteases, currently including mainly eukaryotic proteases, such as mammalian enzymes. They function as proprotein convertases, which are proteases which convert inactive precursor proproteins, including prohormones, progrowth factors, proneurotrophic factors, hormonal receptors, adhesion molecules, retroviral surface glycoproteins, proenzymes and certain protoxins into their active counterparts (Barr, 1991). Proprotein convert-

The classification of subtilisin-like proteases

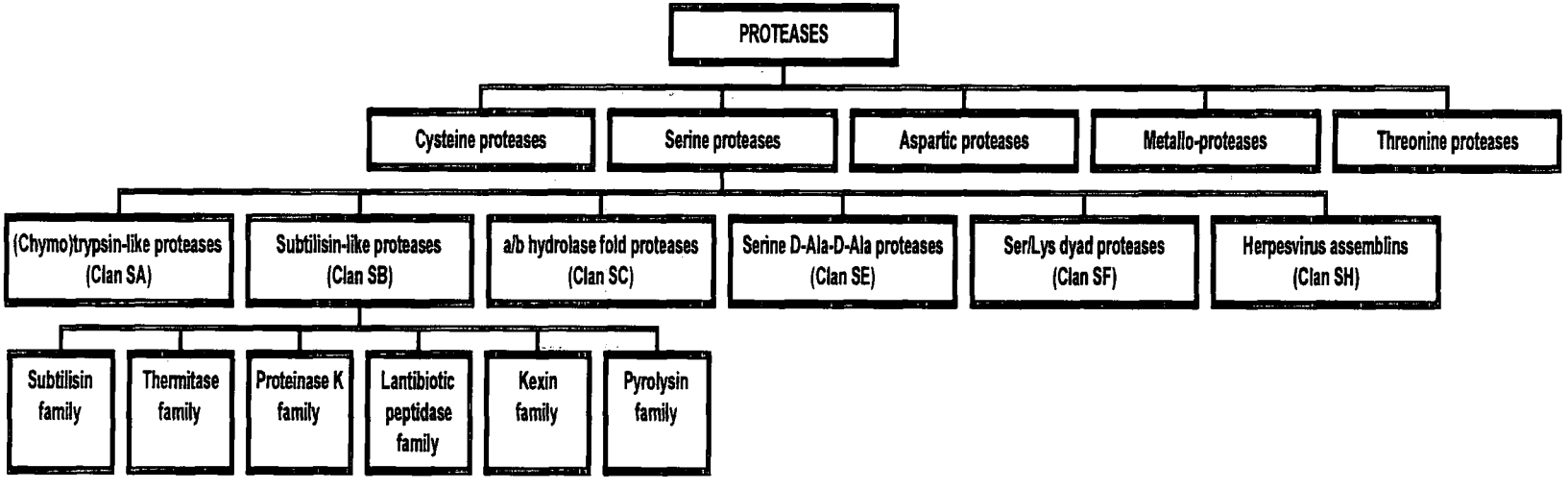


Figure 1.2 A hierarchy of classification showing the different catalytic classes of proteases, the clans of the serine proteases (Rawlings and Barrett, 1994) and the families of the subtilisin-like proteases (Siezen and Leunissen, 1997).

ases process these inactive precursors by proteolytically removing their proregions at sites containing dibasic or multiple basic residues (Barr, 1991).

Members of the pyrolysin family are composed of both proteases and tripeptidases and have been found in bacteria, thermophilic archaebacteria and plants. Most of the plant subtilases examined to date are members of this family. Enzymes in this varied group characteristically contain large insertions and/or long C-terminal extensions, often with associated sequence homology, suggesting a common origin. A further subdivision of the pyrolysins may become apparent as more sequences are discovered (Siezen and Leunissen, 1997). The kexin family of subtilisin-like proteases will be briefly discussed in the next section, before discussing the plant subtilisin-like proteases.

1.4 Subtilisin-like proteases from fungi and animals: the proprotein convertases (PCs)

The discovery of subtilisin-like proteases, with similarity to those found in the *Bacillus* genus, in other bacteria, archaebacteria and eukaryotes has lead to renewed interest in the *in vivo* functions of these proteases. They have been found to have processing functions in lower and higher eukaryotes, having been detected in diverse species, including budding yeast, *Drosophila*, mice, rats and humans (Nakayama, 1997).

The budding yeast gene *KEX2* (*k*iller *ex*pression defective), also called kexin, was discovered due to its involvement in the activation of the pheromone α-mating factor and killer toxin (Leibowitz and Wickner, 1976). It is required for the correct

proteolytic processing of α-factor and killer toxin from larger precursor proteins at sites C-terminal of pairs of basic amino acid residues (Julius *et al.*, 1984). This processing occurs prior to fusion of secretory granules and vesicles with the plasma membrane, and is thought to occur in the Golgi apparatus and associated granules or vesicles (Wilcox and Fuller, 1991). Kexin is a membrane-bound, Ca^{2+} dependent neutral subtilisin-like protease, and is the prototype of a group of eukaryotic proprotein convertases (PCs), which are all involved in the cleavage of important precursor molecules at sites C-terminal to basic amino acid residues (Mizuno *et al.*, 1988; Fuller *et al.*, 1988).

The discovery of kexin led directly to the cloning of homologues from other eukaryotes, by PCR amplification of DNA regions encoding the conserved catalytic sites. Seven distinct PCs have been discovered in mammals: furin, PC2, PC1/PC3, PC4, PACE4 (*p*aired basic *a*mino acid residue *c*leaving *e*nzyme 4), PC5/PC6 and LPC/PC7/PC8/SPC7 (Fuller *et al.*, 1989; see Nakayama,1997 for review). It has been found that proprotein convertases, for example furin, the *fur* (c-*fes/fps* proto-oncogene **u**pstream **r**egion) gene product, can cleave numerous polypeptide hormones C-terminal to pairs of basic residues (such as Arg-Arg, Lys-Arg, Arg-Lys, Lys-Lys), an idea which has been proposed some time ago (Steiner *et al.*, 1967; Chrétien and Li, 1967). The proprotein convertases also process polypeptides after mono-, tri-, tetra- and pentabasic sites, although processing after pairs of basic residues is most common (Barr, 1991).

The proprotein convertases process an extraordinarily wide variety of important precursor proteins and peptides, including most peptide hormones and neuropeptides, growth factors, receptors, adhesion molecules, plasma proteases, matrix metallo-proteases and even bacterial exotoxins. This makes this group of proteases (which

have only been discovered relatively recently) of central importance to the regulation of the biochemistry of eukaryotic organisms.

Furin will cleave many different precursors with the (Arg-X-Arg/Lys-Arg) specificity motif, for instance complement C3, von Willebrand factor (which in humans mediates platelet adhesion and clot formation at sites of vascular injury) and β-nerve growth factor precursors (Nakayama, 1997). Although it is known that a number of prohormones and proproteins are processed by PCs *in vitro*, their endogenous *in vivo* substrates have not been fully defined. The catalytic domains of the different mammalian PCs are highly conserved and tight tissue-specific regulation of these convertases appears to govern the cellular location of proprotein processing.

A number of medical disorders, such as haemophilia and diabetes, can be associated with a malfunction of the normal interaction of PCs with their substrates, due to genetic mutations around the cleavage sites of the precursors. Mutations have been documented in proalbumin, pro-factor IX and insulin pro-receptor (Bentley *et al.*, 1986).

The fact that eukaryotes have PCs has been utilized by viruses. In order to gain entry into host cells, the glycoproteins of the envelope of viruses must be cleaved. PCs are involved in the cleavage of these glycoprotein precursors (for example HIV gp160, measles F_0 and influenza haemagglutinin A glycoproteins). PCs may in fact also determine viral pathogenicity (Horimoto *et al.*, 1994). Depending on the location of the PC(s) which process particular viral components, localized or systemic infections can ensue (Kido *et al.*, 1992). Therefore this affects the different levels of severity of infections which result from exposure to different viruses.

The structure of kexin-like proprotein convertases differs from many other subtilases, in that several other enzyme domains are present, in addition to the well-characterised

subtilisin-like catalytic domain. Signal peptide, propeptide and catalytic domains are usually represented in subtilisin-like proteases in general, however PCs may also contain membrane-binding amphipathic helices, transmembrane domains, cysteine-rich domains, and Homo B regions (Halban and Irminger, 1994). The latter regions and the subtilisin-like catalytic domains are highly conserved among eukaryotic convertases and are required for their normal catalytic activities. Mutation of residues in the Homo B domain can result in loss of convertase activity and also, intriguingly, mis-sorting of the enzyme into the constitutive secretory pathway (Lusson *et al.*, 1997). The precise function of the cysteine-rich region is unclear at present.

The tissue specificities of the PCs have been determined. Furin (Hatsuzawa *et al.*, 1990), PC5/PC6 (Lusson *et al.*, 1993) and LPC/PC7/ PC8/SPC7 (Seidah *et al.*, 1996) seem to have a ubiquitous tissue distribution. PC2, PC1/PC3 (Seidah *et al.*, 1994) and PACE4 (Mains *et al.*, 1997) are expressed at the highest levels in the endocrine and neural tissues. PC4 is expressed exclusively in the germ cells: in testis spermatocytes, spermatids and mature spermatozoa in males, where it may have a role during various stages of spermatogenesis (Nakayama *et al.*, 1992); and in thecal and interstitial cells of the ovary (Mbikay *et al.*, 1997).

At the subcellular level, PCs are generally localized in the *trans*-Golgi network (TGN), and in immature secretory granules and vesicles derived from this compartment (Halban and Irminger, 1994). PC5/PC6 is not concentrated in the TGN, but is found throughout the Golgi complex (Nakayama, 1997). The subcellular location of PC4 has not yet been established. Furin is found in the TGN and also the cell surface, and is thought to cycle between the cell surface and the TGN (Molloy *et al.*, 1994). Generally secretory proteins are transported from their site of synthesis on the rough endoplasmic reticulum to the *cis*-Golgi and then the *trans*-Golgi network via

intercisternal transport. In the TGN, pH-dependent protein sorting occurs and secretion proceeds either by a regulated secretory pathway (where exocytosis is regulated by secretagogues) or a constitutive secretory pathway (where exocytosis occurs at constant rates). PCs operate in both pathways, resulting in regulated and constitutive release of processed proproteins (Halban and Irminger, 1994).

Several PC knockout mice have been generated, enabling the potential *in vivo* roles of PCs to be evaluated and compared to information on the localization of PC mRNAs. Furin (-/-) and PC1/PC3 (-/-) mice were embryo lethal, PC2 (-/-) mice were mildly diabetic and runted, while male PC4 (-/-) mice were infertile (Seidah *et al.*, 1994).

Details about how PCs bring about the maturation of a range of precursors are only now being fully unravelled. It appears that PCs can operate in isolation or in combination with each other to bring about different processing events. This constitutes an extra level of temporal and spatial regulation of bioactive molecule maturation. An example of this is the processing of the precursor molecule POMC (proopiomelanocortin), which can be processed to various bioactive peptides, such as ACTH (adrenocorticotropic hormone), β-LPH (lipotropic hormone), α-MSH (melanocyte stimulating hormone) and β-endorphin. Tissue-specific cleavage of this precursor by PC1/PC3 only, results in the generation of ACTH and β-LPH, but cleavage by PC1/PC3 and PC2, results in α-MSH and β-endorphin production (Seidah *et al.*, 1999b). Thus signalling peptides can be modulated in different ways in different tissues to bring about different downstream effects.

Proprotein convertases have been shown to undergo intramolecular autocatalytic processing in the endoplasmic reticulum (Creemers *et al.*, 1993; Munzer *et al.*, 1997). In fact, the cleaved propeptide of furin acts as a high affinity autoinhibitor of furin until, in acidic and Ca^{2+}-rich conditions (such as those found in the TGN), the

propeptide is further processed, activating furin at its site of action (Anderson *et al.*, 1997).

Recently two mammalian enzymes have been discovered which have highest homology to the pyrolysins. The first of these, S1P (*Site-1 p*rotease), has been implicated in the processing of membrane-bound transcription factors called sterol regulatory element-binding proteins (SREBPs), which are critical in controlling the level of cholesterol and fatty acids in membranes, cells and blood (Sakai *et al.*, 1998). The second of these enzymes, SKI-1 (*s*ubtilisin/*k*exin-*i*sozyme-*1*), has been reported to process brain-derived proneurotrophic factor (Seidah *et al.*, 1999a). The substrate specificity of these pyrolysins differ from the kexin-like PCs in that they exhibit specificity for cleavage C-terminal of the motif (Arg/Lys-X-X-Leu/Thr).

The next section will deal with subtilisin-like proteases which have been found in higher plants.

1.5 Plant subtilisin-like proteases

Cucumisin was the first plant subtilase to be purified (Kaneda and Tominaga, 1975). This enzyme was found in the sarcocarp of developing *Cucumis melo* (musk melon) fruits. A number of different subtilases have been purified from various other cucurbitaceous plants, including *Benincasa hispida* (white gourd; Kaneda and Tominaga, 1977; Uchikoba *et al.*, 1998), *Trichosanthes cucumeroides* (snake gourd; Kaneda *et al.*, 1986), *Cucurbita ficifolia* (fig-leaf gourd; Curotto *et al.*, 1989), *Trichosanthes kirilowii* (yellow snake gourd; Uchikoba *et al.*, 1990) and *Trichosanthes bracteata* (Kaneda and Uchikoba, 1994). These enzymes have been isolated with a

view to using them in the food industry, for example in milk clotting in the dairy industry (Uchikoba and Kaneda, 1996). In the past, most of the plant proteases which have been purified have been shown to be cysteine proteases or, more infrequently, aspartic proteases (Hiraiwa *et al.*, 1997). There are difficulties in the efficient use of cysteine proteases in the food industry, because these enzymes tend to be inhibited by oxidation or metal ions. In order to retain most of their proteolytic activity cysteine proteases, such as papain and ficain, are normally supplied with reducing and chelating agents, but serine proteases, such as cucumisin, do not require these additives (Yonezawa *et al.*, 1997).

Subtilisin-like proteases have also been purified from a variety of other higher plants, such as *Lycopersicon esculentum* (tomato; Vera and Conjero, 1988; Tornero *et al.*, 1996b, 1997); *Helianthus annuus* (the common sunflower; Rudenskaya *et al.*, 1987), *Maclura pomifera* (Osage orange; Rudenskaya *et al.*, 1995), *Lilium longiflorum* (trumpet lily; Taylor *et al.*, 1997) and *Taraxacum officinale* (dandelion; Rudenskaya *et al.*, 1998).

All these proteases belong to the pyrolysin subfamily of subtilases (Siezen and Leunissen, 1997). In general, purified plant pyrolysins have alkaline or neutral pH optima, are relatively thermostable and are inhibited by serine protease inhibitors, for example DFP and PMSF (Bogacheva, 1999).

A novel tomato subtilase, *Le*SBT1, has recently been overexpressed in insect cells and purified (Janzik *et al.*, 2000). Unusually for a plant subtilase, this protease has been shown to have an acidic pH optimum. Acidic conditions promote autocatalytic removal of an amino-terminal inhibitory peptide, which limits its proteolytic activity.

Prodomains of some proteases have been shown to act as selective high affinity inhibitors of the corresponding mature enzymes (Taylor *et al.*, 1995). The proprotein

domains of many proteases can also function as chaperones, which appear to direct the correct folding of the mature protease. This phenomenon was first studied using bacterial prosubtilisin E (Ikemura *et al.*, 1987) and protease prosequences which act in this way have been termed intramolecular chaperones (Shinde and Inouye, 1993). Inactive mature subtilisin seems to fold into an active conformation *in vitro*, even on addition of either prosubtilisin E (Zhu *et al.*, 1989) or just the propeptide domain (Ohta *et al.*, 1991) by an intermolecular mechanism. Unlike the well-studied molecular chaperones, such as the chaperonins and heat shock 70 proteins (hsp70), this process appears to be assisted in a direct and steric fashion by the prosequences of some proteases (Ellis, 1998).

The molecular masses found (and predicted from cDNA sequences) for the mature plant subtilases range from 25kDa, for a serine protease isolated from senescent sunflower leaves (Rudenskaya *et al.*, 1987), to 82kDa for LIM9 found in trumpet lily (Taylor *et al.*, 1997), although the molecular mass of most enzymes was found to lie between 50 and 70kDa.

Several plant subtilisin-like proteases have been examined to determine whether they are glycosylated. All of the proteases investigated, including cucumisin (Kaneda and Tominago, 1975), P69-A (Tornero *et al.*, 1996b), and macluralisin (Rudenskaya *et al.*, 1995), have proved to be glycoproteins. The carbohydrate component of macluralisin makes up over 25% of its molecular weight. No detailed studies into the potential functional significance of the carbohydrate moieties of plant subtilases have been reported.

The first nucleotide sequence encoding a plant subtilase has been reported from a melon cDNA library and has been predicted to encode cucumisin (Yamagata *et al.*, 1994). A pathway for the processing of cucumisin has been presented by the same

authors. The cucumisin gene transcript is thought to be translated as a preproenzyme mainly in the locular tissue around the seeds of the melon fruit. The presequence is a putative signal peptide and the prosequence is thought to act as a molecular chaperone and a specific autoinhibitor to pevent degradation during subsequent secretion of the enzyme (Taylor *et al.*, 1995). All the plant subtilases which have been studied appear to have this modular structure. Cucumisin protease activity is predominantly found in the juice of the central part of the fruits during early development of the fruit (Yamagata *et al.*, 1989). Cucumisin is apparently secreted into this part of the fruit, presumably by initially being translocated into the lumen of the endoplasmic reticulum (ER) through the ER membrane. This is achieved via the signal peptide with concomitant cleavage of this peptide. After secretion of the proprotein into melon juice the protein is believed to undergo autocatalytic processing, removing the prosequence from the rest of the enzyme. This mature protein is proteolytically active and can further proteolytically process itself removing a 14kDa fragment at the carboxyl terminus of the enzyme during the later stages of fruit development (Yamagata *et al.*, 1994). The carboxyl terminal sequence does not appear to be required for protease activity, however it may be essential for secretion across the outer membrane of the cell (Terada *et al.*, 1990) or it may even be involved in an undetermined signalling pathway.

A full-length tomato subtilase cDNA, *Le*SBT1, has been overexpressed in a baculovirus/insect cell system (Janzik *et al.*, 2000). The processing of the overexpressed protein has been studied and found to have similarities to the processing of cucumisin *in planta*. The removal of a putative signal peptide and a prodomain corresponded to these previous observations, however in addition an amino terminal peptide is removed from the mature enzyme. This autocatalytic event occurs at acidic

pH and is necessary for activation of the inactive zymogen (Janzik *et al.*, 2000). It is uncertain whether *Le*SBT1 is further processed by autolytic activity in the plant apoplast like cucumisin. Nevertheless, pH-controlled zymogen activation has been observed amongst the mammalian PCs. Furin (Molloy *et al.*, 1994) and PC2 (Shennan *et al.*, 1995) for example have also been shown to be activated at pH6.0 and below. It may of course be important not to activate these proteases in the plant until they have reached their acidic compartments (such as the apoplast) or until the pH of the compartment reaches the correct level (as occurs in the apoplast during acid growth).

Plant subtilases characteristically degrade various protein substrates such as casein, haemoglobin and collagen. The oxidized B-chain of bovine insulin and artificial peptide substrates have been used to investigate the precise substrate specificities of plant subtilases. These enzymes tend to hydrolyze typical subtilisin substrates, such as Glp-Ala-Ala-Leu-pNA and Glp-Ala-Ala-Phe-Phe-pNA, but not typical trypsin or chymotrypsin substrates (Uchikoba *et al.*, 1995; Bogacheva, 1998). Cucumisin appears to prefer amino acid residues with a small side chain (such as serine, alanine and glycine residues) at the P1' position of the substrate peptide (Arima *et al.*, 2000). The P1' position is the location of the amino acid residue immediately C-terminal of the scissile bond, according to the accepted nomenclature system introduced to describe protease-substrate interactions (Schechter and Berger, 1967). The k_{cat}/K_m values calculated for different synthetic substrates have indicated that the preferred amino acid residues in the P1 position (immediately N-terminal of the scissile bond) include leucine, alanine, phenylalanine and glutamate residues (Arima *et al.*, 2000). The substrate specificity of cucumisin seems to be broad with a preference of hydrophobic amino acids at the P1 position. Peptide hydrolysis by macluralisin and taraxalisin is

very like that displayed by cucumisin (Rudenskaya *et al.*, 1995; Bogacheva *et al.*, 1999).

Similar conclusions can be drawn from substrate specificity experiments using oxidized B-chain of bovine insulin as a substrate. The plant subtilases tested have shown little in the way of obvious preferred cleavage sites, although there appears to be a slight preference towards cleavage of peptide bonds formed by hydrophobic amino acid residues, where these residues occupy the S1 binding site of the enzyme (Bogacheva, 1998).

The endogenous *in vivo* substrates of these enzymes, in so far as there are distinct substrates, have largely remained elusive. Only two endogenous *in vivo* substrates have been suggested for plant subtilisin-like proteases. One of these is the plant hormone systemin, which is discussed below and the other is a leucine-rich repeat protein, which is discussed in section 1.6.

Although all plant subtilases which have been characterized have been shown to belong to the pyrolysin subfamily of subtilases (Siezen and Leunissen, 1997), a number of lines of evidence strongly suggest that kexin-like protease activity also exists in plants (Schaller and Ryan, 1994; Jiang and Rogers, 1999), although currently no such proteases have been purified. In the first study of this type, a 50kDa protein (SBP50) has been identified in plasma membrane preparations, generated from tomato leaves, which binds a biotinylated derivative of systemin (Schaller and Ryan, 1994). Systemin is a plant peptide hormone involved in signalling in response to plant wounding. It has been previously proposed that systemin could be involved in the systemic induction of protease inhibitor genes in tomato in response to wounding or mechanical damage, such as that experienced during herbivory (Pearce *et al.*, 1993). The presence of systemin can induce these genes at levels of femtomoles per young

tomato plant, making it one of the most powerful inducers of its kind (Schaller and Ryan, 1994). The authors of the study found that the amino acid residues of systemin which transpired to be important for binding SBP50, included a furin cleavage site. Systemin is processed *in vitro* at this putative site by a membrane-associated proteolytic activity, which, like furin, can be inhibited by $ZnCl_2$. An antiserum against a *Drosophila* kexin-like protease, dKLIP-1, has been shown to interfere in the binding of biotinylated systemin to SBP50. The antiserum also recognizes a single protein of approximately 60kDa in tomato leaf plasma membrane extracts. These lines of evidence suggest that a membrane-associated kexin-like protease is found in tomato plants.

A second study concerning possible plant kexin-like proteases has been conducted using tobacco suspension culture cells and a series of chimaeric reporter proteins containing kexin-like protease cleavage sites (Jiang and Rogers, 1999). On the basis of the reporter protein cleavage patterns, this investigation has also concluded that a proteolytic activity showing kexin site specificity can be detected, and it appears to be located in the Golgi or TGN (Jiang and Rogers, 1999).

1.6 Proposed functions of plant subtilisin-like proteases

The functions of most of the known plant subtilisin-like proteases, for example macluralisin, Ara12, taraxalisin, cucumisin and the other cucurbitaceous subtilisin-like proteases, have not been determined (see Table 1.2). Where functions have been ascribed to plant subtilisin-like proteases, involvement in these processes is merely suspected or is certainly only very poorly understood.

Subtilase	Plant	Association	Reference
cucumisin	*Cucumis melo*	Unknown	Kaneda and Toninaga, 1975
snake gourd subtilase	*Trichosanthes cucumeroides*	Unknown	Kaneda *et al.*, 1986
P69A	*Lycopersicum esculentum*	Pathogenesis (?)	Vera and Conjero, 1988
macluralisin	*Maclura pomifera*	Unknown	Rudenskaya *et al.*, 1995
Ag12-2	*Alnus glutinosa*	root nodule development	Ribeiro *et al.*, 1995
Ara12	*Arabidopsis thaliana*	Unknown	Ribeiro *et al.*, 1995
TMP	*Lycopersicum esculentum*	Microsporogenesis	Riggs and Horsch, 1995
P69B	*Lycopersicum esculentum*	Pathogenesis	Tornero *et al.*, 1997
LIM9	*Lilium longiflorum*	Microsporogenesis	Taylor *et al.*, 1997
taraxalisin	*Taraxacum officinales*	Unknown	Rudenskaya *et al.*, 1998
AIR3	*Arabidopsis thaliana*	lateral root formation	Neuteboom *et al.*, 1999
P69C P69D	*Lycopersicum esculentum*	Pathogenesis unknown	Jordá *et al.*, 1999
*Le*SBT1, 2, 3, 4(A-E) P69E, F	*Lycopersicum esculentum*	Unknown unknown	Meichtry *et al.*, 1999

Table 1.2 Functional associations of plant subtilases. This is not a comprehensive list as some of the subtilases which have been purified from cucurbitaceous plants have been omitted. However, very little is known about the *in vivo* function of these enzymes. Although the role of the P69A, P69D, P69E and P69F subtilases has not been established yet, all P69 enzymes share close sequence homology; this subgroup may also share functional homology.

Presently fifteen genes are predicted to encode subtilisin-like proteases in tomato: P69A-F, SBT1-3, SBT4A-E and TMP (Meichtry *et al.*, 1999). Two of the corresponding P69 proteases have been purified from tomato leaves which were infected with citrus exocortis viroid (CEV) (Vera and Conjero, 1988; Tornero *et al.*, 1997). P69B (Tornero *et al.*, 1997) and P69C (Jordá *et al.*, 1999) expression is induced in tomato leaves and stems in response to this viroid and other pathogens (Tornero *et al.*, 1996a) and therefore they are classed as pathogenesis-related (PR) proteins (Granell *et al.*,1987). Some PR proteins have been demonstrated to be directly involved in an induced defence response by the plant against invasive pathogens (Bowles, 1990).

The different tomato subtilase genes show a complex pattern of tissue-specific expression (Meichtry *et al.*, 1999), for example P69E and P69F transcripts are primarily expressed in roots and hydathodes, respectively (Jordá *et al.*, 2000). The genes encoding the four proteases P69A-D are located together in a genomic cluster (Jordá *et al.*, 1999). P69A and P69D are expressed constitutively, showing different patterns of expression throughout development. As mentioned previously P69B and P69C are greatly upregulated as a result of infection with pathogens, such as CEV or *Pseudomonas syringae*, or spraying with salicylic acid (Jordá *et al.*, 1999). The primary sequences of proteases P69A-F are all extremely similar, as are those of SBT3 and SBT4A-E (Meichtry *et al.*, 1999). Members within these two groups may have similar roles to play in the plant even though they appear to be differentially regulated. Apart from putative roles in pathogenesis for individual P69 proteases and microsporogenesis for TMP, the precise function of the various different tomato subtilases still remains uncertain.

In addition to systemin, evidence has been obtained for just one other possible *in vivo* substrate of a plant subtilisin-like protease. The DNA encoding this proposed protein substrate has been identified in tomato and its mRNA is slightly upregulated in response to CEV infection (Tornero *et al.*, 1996a). It was isolated by differentially screening a cDNA library, constructed from CEV-infected tomato plants, using cDNA synthesized from healthy and CEV-infected leaf tissue (Tornero *et al.*, 1996a). The cDNA encodes a leucine-rich repeat protein (LRP) containing a potential leucine zipper motif (Landschulz *et al.*, 1988) and another repeated leucine-rich motif. LRP-specific antisera have been used to study the presence of LRP in crude extracts of healthy and CEV-infected tomato plant tissue. In healthy tissues one immunoreactive protein of 35kDa (assumed to correspond to LRP) has been observed. In CEV-infected tissues, LRP is believed to be proteolytically processed as two immunoreactive variants of 33kDa and 35kDa (Tornero *et al.*, 1996a). Purified P69A subtilisin-like protease has been incubated with crude extracts of healthy tissue at ratios of protease to total protein corresponding to those seen in viroid-infected leaves. *In vitro* processing of LRP by P69A has been shown by Western blotting of these samples. The N-terminus of LRP is thought to be processed, possibly at a site adjacent to two arginine residues. It remains to be seen whether the *in vitro* processing of LRP by a P69 subtilisin-like protease resembles *in vivo* events occurring during pathogenesis. It is also unknown whether LRP constitutes a specific target for this protease, or whether a number of different proteins are processed by the enzyme.

Leucine-rich repeat proteins are known to be involved in protein-protein interactions and frequently participate in signal transduction (Kobe and Deisenhofer, 1994). Proteins containing leucine-rich repeat (LRR) motifs fall into several distinct categories on the basis of their established functions. Signal-transducing receptors

containing these motifs, such as CD14 and the protooncogene Trk (Schneider and Schweiger, 1991), are found at the cell surface and trigger the phosphorylation of intracellular proteins in response to ligand binding. Analogous transmembrane receptor protein kinases are found in plants. For instance, in *Arabidopsis* the TMK1 (Chang *et al.*, 1992) and RLK5 (Walker, 1993) proteins contain LRR motifs thought to be found in the extracellular matrix. Other proteins containing LRR motifs include secreted cell adhesion molecules (CAMs) found in animals. CAMs are important in controlling morphogenesis. Examples of these include the human glycoprotein Gp1b (Lopez *et al.*, 1987) and the *Drosophila* protein Toll, with its well-defined role in embryogenesis (Keith and Gay, 1990). Despite their prevalence in the extracellular matrix, proteins containing the LRR motif also occur intracellularly, such as the disease-resistant gene product of the RPS2 gene in *Arabidopsis* (Bent *et al.*, 1994). Polygalacturonase inhibitor proteins (PGIPs) are also rich in leucine residues, for example 15% of tomato PGIP consists of leucine residues (Stotz *et al.*, 1994).

Local cell-cell communication is an important facet of animal development (Bronner-Fraser, 1990), but it is poorly understood in plants (Dixon and Lamb, 1990). In animals proteases are thought to process protein substrates in the extracellular matrix as part of signal transduction events associated with morphogenesis, tissue repair, wound healing and in disease states (Alexander and Werb, 1991; Chen, 1992). The demonstration that LRP is processed *in vitro* by a P69 subtilisin-like protease could indicate that a similar interaction between proteases and extracellular proteins occurs in plants.

A different subtilisin-like protease, which appears to be meiosis-specific has been purified from *Lilium longiflorum* (trumpet lily) microsporocytes (Taylor *et al.*, 1997), and from N-terminal sequence data, the protease was identified as LIM9 (*li*ly *i*nduced

at *m*eiosis 9). The cDNA encoding LIM9 has previously been isolated by a subtractive hybridisation approach (Kobayashi *et al.*, 1994). LIM9 is expressed principally in anther locules, specifically in the single layer of tapetal cells which surround the microsporocytes, but lower levels of expression are also seen in the microsporocytes themselves. Tapetal cells are thought to act as nurse cells for the developing microsporocytes. Connections between tapetal cells and microsporocytes have been observed during early development (Spitzer, 1970). The exact role of LIM9 has not been determined, but several possibilities have been proposed (Taylor *et al.*, 1997). This serine protease may be involved in the degradation of the callose wall which envelops the microspores. This process could be aided by LIM9 if it generates gaps in the primary cell wall of the microspores, enabling callase to penetrate and digest the callose wall. LIM9 could take part in the apoptosis of tapetal cells, an event necessary for anther dehiscence (Goldberg *et al.*, 1993). Alternatively, LIM9 may function in the nutrition of the developing microsporocytes, which may explain the need for extensive connections between tapetal cells and microsporocytes during development. LIM9 could conceivably function as a proprotein convertase in a manner analogous to the characterized mammalian PCs, and thus have a role in signal transduction events concerning microsporogenesis. A cognate clone, encoding tomato meiotic proteinase (TMP), has been isolated from tomato plants and is also expressed only in young anthers (Riggs and Horsch, 1995). Proposed antisense studies in tomato should help to elucidate the role of these meiosis-associated proteases further.

A cDNA encoding the Ag12 subtilisin-like protease has been isolated from a cDNA library made from *Alnus glutinosa* (alder) root nodule poly (A) RNA (Ribeiro *et al.*, 1995). The tissue expression of *ag12* has been determined by *in situ* hybridization of root nodule sections using ^{35}S-labelled sense and antisense RNAs (Ribeiro *et al.*,

1995). *ag12* is expressed at high levels during the early stages of root nodule development brought about as a result of the symbiotic relationship with the actinomycetes (nitrogen-fixing bacteria) of the genus *Frankia*. Highest levels of *ag12* expression have been visible in young infected cells, which are actively enlarging as a consequence of infection. Lower levels of expression have been seen in the fixation zone of the nodule, where the symbiotic relationship has been properly established. Senescent infected cells of the root nodule did not appear to express *ag12*. Ag12 is thought to be targeted to the extracellular matrix. There its activity may not be limited to hydrolysis of plant proteins as it could potentially process one or more proteins in the bacterial cell wall. Indeed expression of *ag12* may be an antibacterial response. The nature of the involvement of Ag12 in the establishment of a symbiotic relationship in the developing root nodule is far from clear. Northern analysis indicates that *ag12* is primarily expressed in the root nodules, but also, to a much lesser extent, in shoot tips (Ribeiro *et al.*, 1995). There may be a link with both of these plant tissues, as both contain regions which are undergoing rapid extension growth. Ag12 might therefore be involved in localized loosening of the plant cell wall, prior to cell extension, however its role remains to be clarified.

A homologue of *ag12*, known as *ara12*, has been isolated from *Arabidopsis thaliana* (Ribeiro *et al.*, 1995), and *ara12* cDNA encodes a subtilisin-like protease (see Figure1.3). From the predicted amino acid sequences of Ag12 and Ara12 it is apparent that they both share significant homology with prokaryotic subtilisins and other eukaryotic subtilisin-like proteases, particularly around the active site residues (Siezen *et al.*, 1991). The presence of a putative transmembrane signal peptide has been predicted, suggesting that Ara12 protein is targeted to the extracellular matrix (Ribeiro *et al.*, 1995; von Heijne, 1986). Northern analysis has revealed that *ara12* is expressed

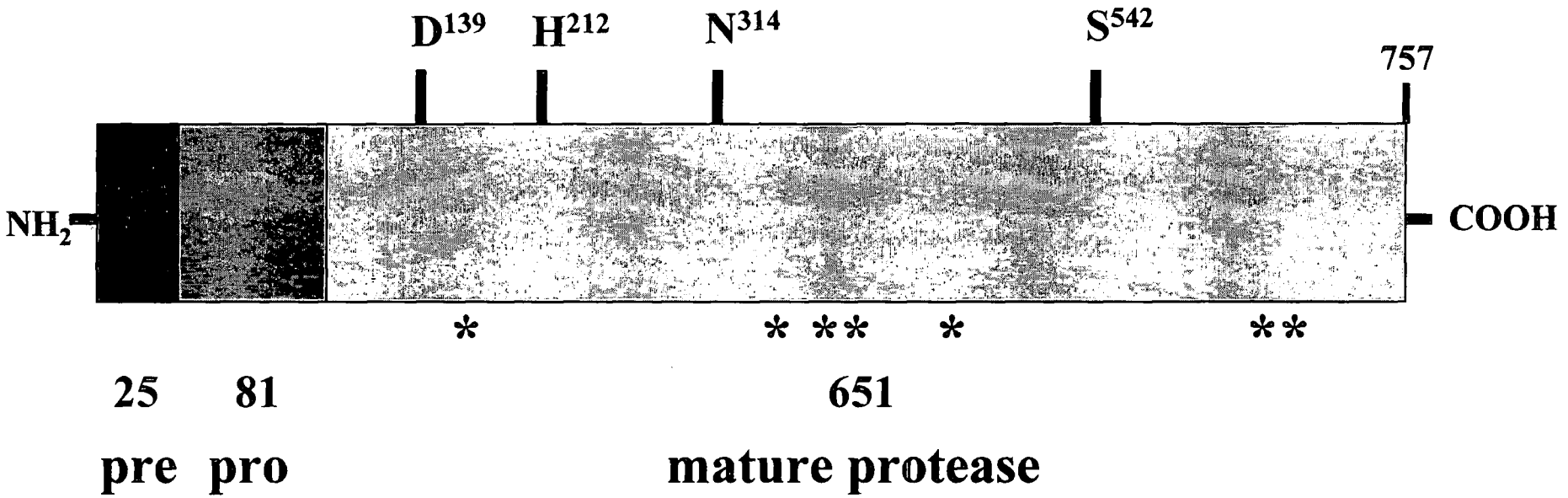


Figure 1.3 Primary structure of prepro-Ara12 subtilisin-like protease from *Arabidopsis thaliana*. The pre-sequence is a putative signal peptide and the proprotein region may be involved in the folding of the mature protease. Amino acid residues predicted to be involved in proteolytic cleavage are shown. Possible glycosylation sites are represented by asterisks.

in all *Arabidopsis* tissues examined and hybridizes to a 2.5kb mRNA transcript (Ribeiro *et al.*, 1995). Highest expression levels have been seen in immature silique tissue, while lower relative levels have been witnessed in flower, leaf and stem tissues. Root tissue showed very low relative levels of *ara12* expression. The authors suggest that Ara12 may be involved in plant development and that, in the case of the *Alnus glutinosa* protein Ag12, a plant subtilase with a role in non-symbiotic development has acquired a specific role in the interaction with symbionts (Ribeiro *et al.*, 1995). It should be stressed that very little is known about Ara12 protease, especially its function.

Another cDNA which has been shown to encode a subtilisin-like protease has been isolated from *Arabidopsis thaliana*. This cDNA, *AIR3* (*AIR* for *a*uxin-*i*nduced in *r*oot cultures), has been obtained by differential screening of an auxin-treated root cDNA library (Neuteboom *et al.*, 1999). Auxins are involved in the regulation of plant growth and development, for example lateral root formation is induced by auxin (Blakely *et al.*, 1988). *AIR3* is also thought to encode an extracellular protein, as the predicted primary sequence of the protease contains a putative signal peptide required for translocation across the plasma membrane. A number of gene products thought to be involved in lateral root formation, such as the structural hydroxyproline-rich glycoproteins (HRGPs) are located in the cell wall. These proteins, also known as the extensins, are thought to form a cross-linked network strengthening the plant cell wall (Lamport and Epstein, 1983). It has been postulated that local modification of the cell wall is a vital step in the initiation of lateral root formation (Neuteboom *et al.*, 1999). Extensins, such as HRGPnt3 from tobacco, have been found which are primarily expressed in the cells destined to form lateral root buds, as well as being present in the tips of emerging lateral roots (Vera *et al.*, 1994). Deposition of these proteins during

root formation probably indicates that they are required to strengthen parts of the cell wall, enabling new growth to occur in adjacent regions. The rigid cell walls of growing plant cells are known to undergo a process of loosening brought about mainly by the action of expansins on the cellulose-hemicellulose network (Cosgrove, 1999). Proteases, such as the subtilisin-like proteases, may have a direct or indirect role in the relaxation of the extensin network, a process which may also have great importance in the growth of cells surrounded by a cell wall consisting of carbohydrates and proteins. Although it is speculated that AIR3 is involved in lateral root formation, its precise role is unknown (Neuteboom *et al.*, 1999).

1.7 The plant extracellular matrix

Plant subtilisin-like proteases have been assumed to be found in the extracellular matrix, in view of the nature of their putative signal peptides. Recent studies with *Arabidopsis* cells have demonstrated an extracellular location for the Ara12 subtilisin-like protease (Robertson *et al.*, 1997). It is worth considering the purpose and the composition of the extracellular matrix in higher plants. Extracellular components influence the architecture of the cell wall and determine the manner in which plant cells grow.

1.7.1 Cell walls and their roles in plants

In plants protoplasts are surrounded by cell walls and intercellular material, together forming the apoplast. A cell wall also surrounds most prokaryotic protoplasts. An enormous amount of evidence has accumulated which shows that the plant cell wall is

best considered as an extracellular matrix with dynamic properties (Bolwell, 1993). In higher plants, formation of the plant cell wall begins during the last stage of cell division. A thin primary cell wall is deposited on each side of the middle lamella of the cell plate. The primary cell wall is composed of approximately 90% polysaccharides and 10% glycoprotein. Components in the primary cell wall include fibrous cellulose running through a matrix of complex polysaccharides, known as hemicelluloses and pectic polysaccharides, (glyco)proteins and free spaces filled with water. For a more detailed list of these components see Table1.3. Plant cells, such as leaf mesophyll cells do not develop a secondary wall, but in many plant cells a secondary wall of a single, or multiple layers is deposited inside the primary wall. Cell walls, in so far as they are found in algae (in Rhodophyta, Chlorophyta, Phaeophyta and Xanthophyceae), tend to contain a greater variety of polysaccharides than is seen in higher plants (Percival and McDowell, 1981). Chitin is a major component of cell walls in fungi, although cellulose can be found in the cell walls of some fungi (Wessels and Sietsma, 1981). The cell walls of bacteria and blue-green algae (prokaryotes) are unique in that they contain mucopeptide.

The cell wall has proved to perform numerous important functions in plants. Both the size and the shape of plant cells is determined by their cell walls. During the development of higher plants there is an increase in the number and the size of cells.

Cell walls provide mechanical strength to plant cells and tissues. Turgor pressure is set up as the aqueous contents of the cells push against the rigid cell walls. Turgidity of cells plays a major role in supplying mechanical strength to plant organs. A cellulose-xyloglucan network provides the cell with high tensile strength: about 70% of the total strength of normal primary cell walls (Shedletzky *et al.*, 1992). Lignin in cell walls

Phase	Components	
microfibrillar	**cellulose (β1,4-glucan)**	
matrix	**Hemicelluloses**	xyloglucan xylan glucomannan galactoglucomannan mannan galactomannan glucuronomannan callose (β1,3-glucan) β1,3-, β1,4-glucan arabinogalactan II
	Pectins	rhamnogalacturonan I arabinan galactan arabinogalactan homogalacturonan apiogalacturonan rhamnogalacturonan II
	Proteins	hydroxyproline-rich glyco-proteins (HPGPs) or extensins glycine-rich proteins (GRPs) proline-rich proteins (PRPs) arabinogalactan proteins (AGPs) solanaceous lectins others, including enzymes
	Phenolics	lignin ferulic acid others, e.g. coumaric acid, truxillic acid
	Others	cutin suberin waxes sporopollenin

Table 1.3 Plant cell wall components. Not all the matrix components listed are found in all plant cell walls. (Table modified from Brett and Waldron, 1996).

confers increased mechanical strength and rigidity, and the ability to withstand compressive forces. This is a particularly important factor in aerial parts of a plant, which must counteract gravity and the wind and rain. Depositing silica into the walls of grasses is also an effective means of bringing about support.

Plant cell walls also act as physical barriers, protecting the protoplasm of the organism from its environment, restricting the diffusion of solutes and preventing the invasion of the plant by viruses, bacteria and fungi, or attack by insects. Various intricate mechanisms for the detection of plant pathogens have been recognized. Cell wall fragments have been shown to elicit plant responses concerned with the regulation of growth and development and defence against pathogen attack (Bolwell, 1993). Some wall fragments originating from pectic polysaccharides can induce the production of antimicrobial compounds (phytoalexins). Elicitors, such as, certain oligo-galacturonides from the walls of plants (and β-glucans from fungal cell walls), are known to result in the induction of phytoalexin synthesis in plant tissue (Lawton and Lamb, 1987). Thus a plant pathogen invading the cell wall and deploying its cell wall-degrading enzymes, releases oligosaccharides from the wall, and this elicits the production of antimicrobial compounds at the site of infection.

Fragments derived from pectic polysaccharides are also capable of stimulating systemic synthesis of insect and microbial proteinase enzyme inhibitors, in response to mechanical damage (Ryan and Farmer, 1991). It has been proposed that a PIIF (proteinase inhibitor inducing factor) acts as a systemic signal in the plant to induce the production of protease inhibitors. Furthermore, it is believed that these protease inhibitors protect the plants by reducing the efficacy of the digestive enzymes of the invading pathogen (Ryan and Farmer, 1991).

Fragments of plant cell wall polysaccharides are known to elicit a range of regulatory activities in growth and development within the plant. Oligosaccharides with such functions have been termed "oligosaccharins" (Albersheim *et al.*, 1993). Different oligosaccharin fragments from the primary cell wall of suspension-cultured sycamore cells are capable of inducing hypersensitive cell death (Yamazaki *et al.*, 1983), inhibition or stimulation of 2,4-D (2,4-dichlorophenoxyacetic acid) induced growth of pea epicotyls (York *et al.*, 1984) and inhibition of flowering and induction of vegetative growth in *Lemna gibba* (duckweed; McNeil *et al.*, 1984). Other oligosaccharin activities include the inhibition of root growth, lignification, activation of ion transport, depolarization of the plasma membrane, stimulation of an oxidative burst and the promotion of fruit ripening (Fry *et al.*, 1993).

Primary cell walls are reasonably porous to allow the passage of water and low molecular weight solutes (Carpita *et al.*, 1979). As cells differentiate into specialized tissue lignification (in fibres, vessels and tracheids), cuticularization (in epidermal cells) and suberization (for example, in cork cells) of the wall leads to a reduction in wall permeability. It is possible that the size of pores is under the fine control of pectic polysaccharides, which could thereby control access of large molecules such as enzymes to their substrates. This may be one way in which cell wall metabolism is regulated.

Undoubtedly the complexity and entire range of the functions of the plant cell wall has not yet been fully understood. It is important to study and attempt to comprehend the biology of plant cell walls, especially as these structures are such a vital source of food and dietary fibre.

The next section deals with the proteins found in the plant extracellular matrix. The proteolytic activity of subtilisin-like proteases is expected to be directed against this

component of the plant extracellular matrix in ways which have not been fully defined yet.

1.7.2 Proteins and glycoproteins in the plant cell wall

Structural proteins and enzymes are an important and integral part of the plant cell wall (see Table1.4). Structural proteins in the plant cell wall fall into five main classes: the hydroxyproline-rich glycoproteins (HRGPs), also known as the extensins, the glycine-rich proteins (GRPs), the proline-rich proteins (PRPs), the solanaceous lectins and the arabinogalactan proteins (AGPs).

The word "extensin" was coined by Lamport (1963) to refer to HRGPs and their presumptive role in cell wall extensibility. Extensins are found in the the cell walls of higher plants, being particularly abundant in dicots. These basic glycoproteins consist of 50% protein and are rich in 4-hydroxyproline, serine, valine, threonine, lysine, tyrosine and an unusual tyrosine derivative, isodityrosine. It is thought that isodityrosine may be synthesized *in vivo* through a wall peroxidase-mediated oxidative coupling of tyrosyl residues (Fry, 1986). The presence of intramolecular isodityrosine cross-links in extensins probably explains their remarkable insolubility. This insolubility led to a theory that extensins are covalently cross-linked in the cell wall. It has also subsequently been suggested that extensins are covalently cross-linked to pectin (Showalter, 1993). Extensins may also interact ionically with pectic polysaccharides on account of their positively charged residues. If so, changes in pH or Ca^{2+} concentration could regulate those interactions. Extensins often contain the repeating pentapeptide motif, Ser-Hyp_4, sometimes within larger repeating motifs. The oligosaccharide component of extensins consists mainly of arabinosyl residues and a

Function	Protein
structural proteins	hydroxyproline-rich glycoproteins (HRGPs) or extensins glycine-rich proteins (GRPs) proline-rich proteins (PRPs) arabinogalactan proteins (AGPs) solanaceous lectins gumarabic proteins
carbohydrate modifying enymes **a) glycosidases (hydrolases)**	glucosidase (glucanase/cellulase) galactosidase (galactanase) mannosidase (mannase) arabinosidase (arabinase) trehalase xylosidase (xylanase) galacturonosidase (pectinase; polygalacturonase) glucuronosidase fucosidase cellobiohydrolase *N*-acetyl glucosaminidase (chitinase) fructofuranosidase (invertase) pectin methyl esterase pectin acetyl esterase feruloyl esterase *p*-coumaroyl esterase pectin/pectate lyase (sucrose hydrolases, galactosides, glucosides, etc.)
b) glycosyl transferases	xyloglucan endotransglycosidase (XET)
other enzymes	Peroxidases oxidases (eg. Ascorbate-, malate-, polyamine-, polyphenol-; laccases) dehydrogenases (eg. malate-, pyruvate-) proteases (eg. Serine-) phosphatases ribonucleases
miscellaneous proteins	Expansins defensins cysteine-rich thionins water-regulated proteins histidine-tryptophan-rich protein

Table 1.4 Proteins and glycoproteins found or expected in plant cell walls.

smaller quantity of galactosyl residues.

Extensins form the major structural protein component of the cell wall, and may have a function in cell expansion, wound healing and plant defence. Extensin is tightly cross-linked and it may be possible that peroxidase, by catalyzing the formation of isodityrosyl, tightens the cell wall further and limits cell expansion. Intriguingly it has been shown that soluble extensin monomers can form an oligmer in the presence of a cell wall extract (Everdeen *et al.*, 1988). This effect is enhanced through physiological stress or the application of elicitor or abiotic elicitors, such as glutathione. This process is believed to be catalyzed by peroxidase and results in an "ultrarapid" defence mechanism (Bradley *et al.*, 1992). Extensin in the cell wall has also been proposed to agglutinate and immobilize plant pathogens on account of its polycationic nature.

The primary structure of GRPs contains up to 70% glycine organized into short amino acid repeat units. At least two classes of GRPs exist. One class is specifically localized in primary xylem cell walls (Ye and Varner, 1991) and is developmentally regulated. The second class is localized in the cytoplasm and is regulated by stress conditions. Cell wall GRPs presumably have important functions in vascular development and wound healing.

The PRPs typically contain Pro-Pro repeats within various other motifs. They contain equimolar quantities of proline and hydroxyproline and appear to be only lightly glycosylated. PRPs consist of two classes: components of plant cell walls and plant nodulins (Jose and Puigdomenech, 1993). Like extensins, PRPs are thought to be insolubilized in the wall and may interact ionically with other cell wall components, such as the acidic pectins. It seems most probable that PRPs are involved in nodule morphogenesis and signalling in the bacterial infection process.

Lectins are proteins, or glycoproteins, which are able to bind non-covalently to carbohydrates. Solanaceous lectins are localized in the walls of solanaceous plants, such as potato and tobacco plants. Potato tuber lectin (PTL) consists of 50% protein, and is particularly rich in hydroxyproline, serine, glycine and cysteine. Its carbohydrate component (50%) contains mainly arabinosyl residues linked to hydroxyproline. The function of plant lectins is unknown, although roles in cell-cell interaction have been suggested, for example in sugar transport, stabilization of seed storage proteins and control of cell division.

Arabinogalactan proteins resemble HRGPs, although the protein moiety accounts for just 2-10% of AGPs. These glycoproteins are acidic, rich in hydroxyproline, serine, alanine and glycine. They are extensively glycosylated and are very soluble. They remain highly soluble in the cell wall, and have been proposed to act as glues, lubricants and humectants (Fincher *et al.*, 1983). AGPs may function in wound healing or plant defence, however no firm evidence has been put forward to support this.

A variety of different enzymes are located in plant cell walls catalyzing a host of separate reactions (Table1.4 lists these and other cell wall proteins). Usually plant cell wall enzymes are glycoproteins. Certain enzymes are involved in the biosynthesis of cell wall polymers, for example peroxidase and NADH malate dehydrogenase. It is thought that wall-bound peroxidase cross-links extensin via isodityrosine and/or pectic fragments via diferuloyl bridges (Fry, 1986). Peroxidases are important in the defence mechanisms of the plants. Also peroxidases and NADH malate dehydrogenase are required in the oxidative polymerization of lignin monomers. A large group of wall-associated enzymes, the polysaccharide hydrolases also participate in defence, as well as governing changes in the structure of the cell wall (such as during extension growth or abscission) and changes during differentiation (such as xylem vessel

differentiation). Examples of these cell wall hydrolases include: β-1,3-glucanases, cellulase, arabinosidases, β-fructofuranosidases, α- and β-galactosidases, α-, and β-mannosidases, trehalases, β-glucosidase, β-glucuronidase, β-xylosidases and acid phosphatases (Bowles, 1990). Other enzymes which are present in the cell wall are involved in nutrient modification, before being allowed to pass into the protoplast. Examples of these cell wall enzymes include: hydrolases for sucrose, β-galactosides, β -glucosides, sugar phosphate esters and proteases (Fry, 1995). Other cell wall proteins, which are non-enzymic can be found, for example cysteine-rich thionins, water-regulated proteins and a histidine-tryptophan-rich protein (Florack and Stiekema, 1994).

1.7.3 Macromolecular interactions and cell wall architecture

Plant cell walls are extremely complex laminate structures and receive their characteristic properties and architecture through a series of highly cross-linked (Fry, 1986; Kieliszewski and Lamport, 1994) polymer networks composed of cellulose-hemicellulose (Carpita, 1983; McCann and Roberts, 1991), pectin (Morris *et al.*, 1982; Shedletzky *et al.*, 1992), extensin (Lamport and Epstein, 1983; Fry, 1986) and lignin (Whitmore, 1981; Meyer *et al.*, 1991). Precise models of cell wall architecture are still under debate, particularly the relative importance of the cellulose-xyloglucan network (Talbott and Ray, 1992). Much work is required to establish a model of the cell wall which can do justice to the manner in which all the components interact. A model of the primary cell wall of dicotyledonous plants (Brett and Waldron, 1996) is shown in Figure 1.4. The cross-links thought to occur in plant cell walls (Fry, 1986) are illustrated in Figure 1.5.

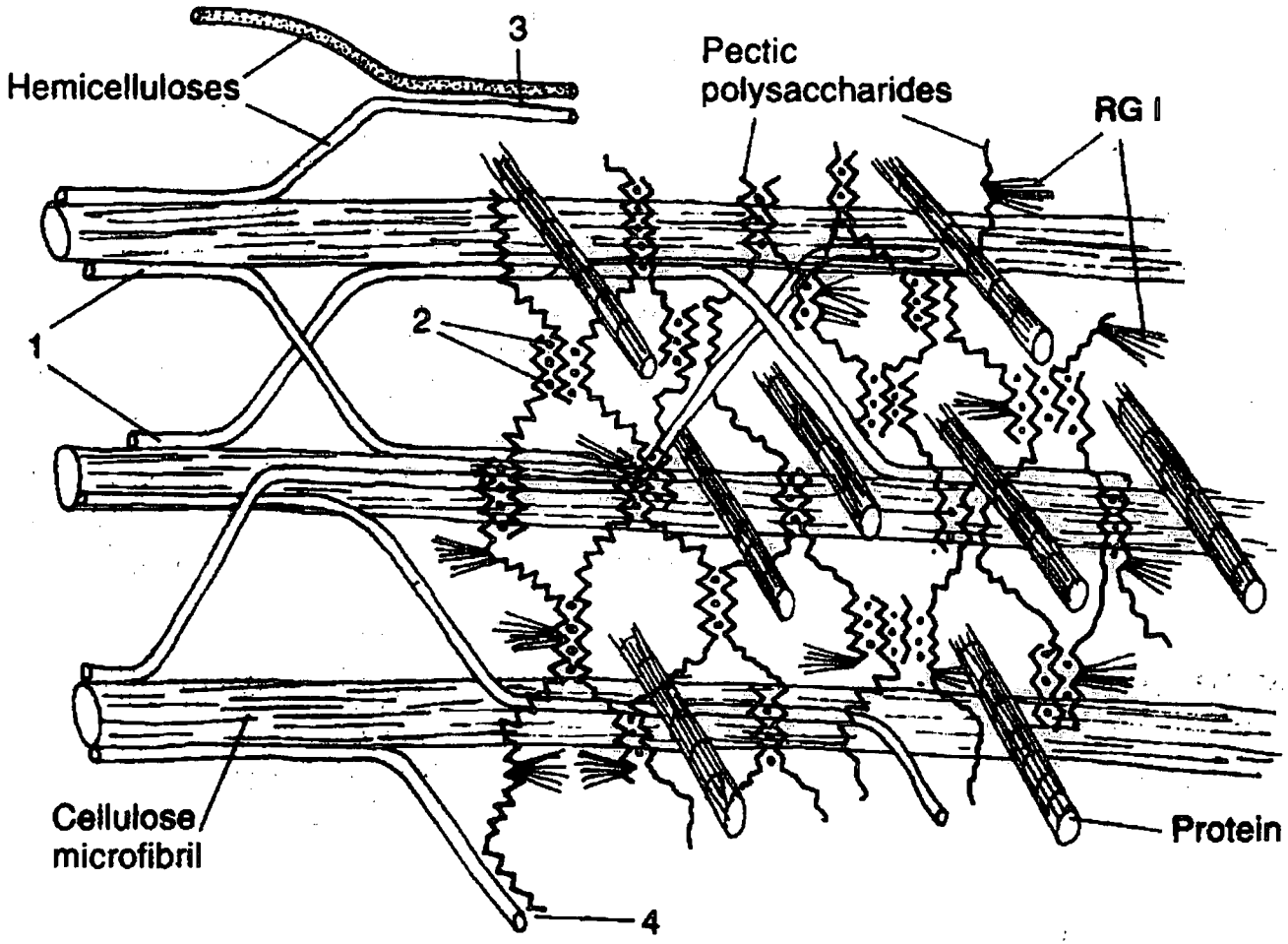


Figure 1.4 Schematic representation of a primary cell wall of a dicotyledonous plant, showing possible interactions between several classes of cell wall polymer and their spatial arrangement. This representation is a much simplified version of a primary cell wall. Interactions between cellulose: hemicellulose (1), pectin: pectin (2), hemicellulose: hemicellulose (3) and hemicellulose: pectin (4) are shown. RG I = rhamnogalacturonan I. Figure taken from Brett and Waldron, 1996.

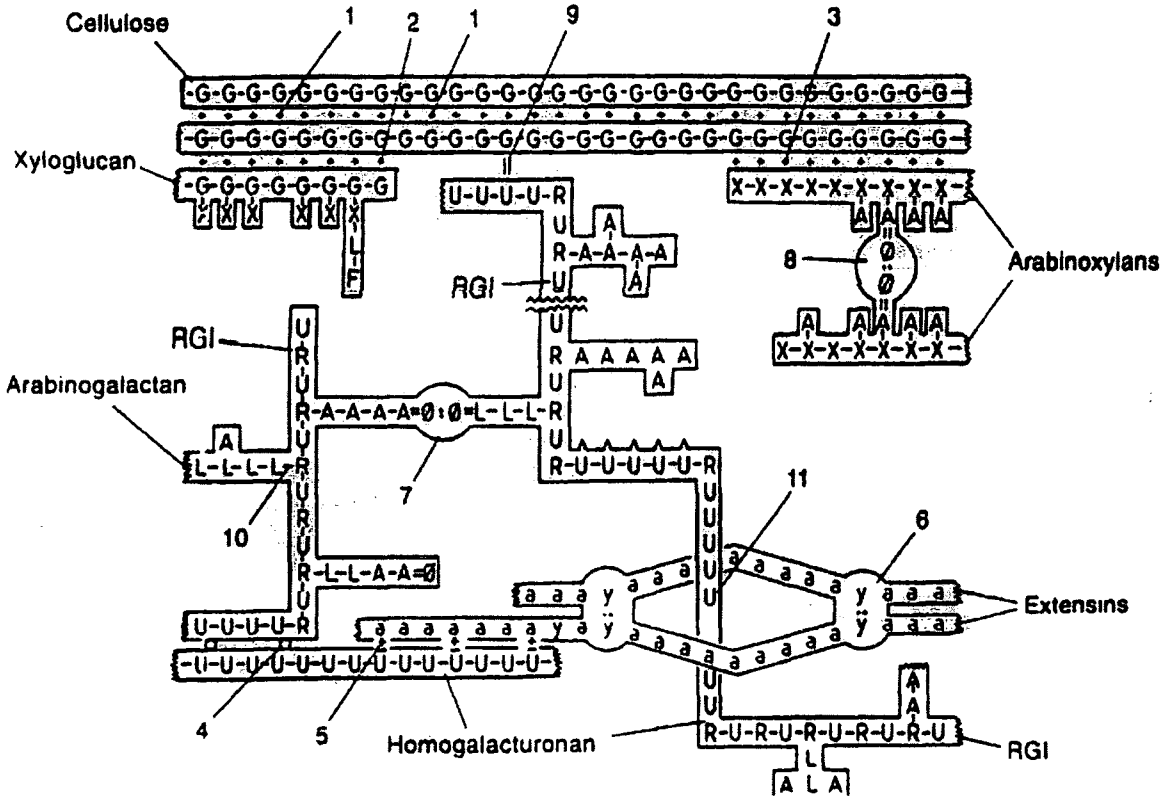


Figure 1.5 Representative primary structures and possible cross-links of wall polymers. The numbers below on the left refer to the interactions between different types of cell wall polymers illustrated above. The symbols on the right refer to the nature of those interactions. The diagram does not constitute a model of a plant cell wall and no significance is placed on the size and relative spacing of the molecules portrayed. (Figure adapted from Fry, 1986).

1. cellulose-cellulose	(+) hydrogen bonds
2. xyloglucan-cellulose	
3. xylan-cellulose	
4. homogalacturonan-homogalacturonan	(o) calcium bridges
5. extensin-pectin	(±) other ionic bonds
6. extensin-extensin	(:) coupled phenols
7. pectin-pectin	
8. arabinoxylan-arabinoxylan	
9. pectin-cellulose	(=) ester bonds
10. arabinogalactan-rhamnogalacturonan	(-) glycosidic bonds
11. pectin-in-extensin	(#) entanglement

Abbreviations: A, arabinose; F, fucose; G, glucose; L, galactose; R, rhamnose; RGI, rhamnogalacturonan I; U, galacturonic acid; Û, galacturonic acid methyl ester; a, amino acid (other than tyrosine); y, tyrosine; y:y, isodityrosine; ∅, ferulic acid; ∅:∅, diferulic acid.

Cell walls are selectively extended during the growth of plant cells. A number of different agents have been implicated in regulating the relaxation of the cell wall, a prerequisite for cell expansion (Taiz, 1984). Plant cell growth is stimulated by protein factors particularly under acidic conditions, which have been shown to be induced by auxin (Rayle and Cleland, 1992). A group of proteins called expansins appear to be the principal mediators of this acid growth (Cosgrove, 1993; Shcherban *et al.*, 1995). Expansins are thought to bring about relaxation of the plant cell wall by binding to cell wall polymers and destabilizing glycan-glycan hydrogen-bonding (McQueen-Mason and Cosgrove, 1994).

Hydrolytic enzymes in the cell wall may also cause loosening of the wall by hydrolyzing bonds in the matrix, for example glucosidic bonds in matrix polysaccharides are cleaved by endoglucanases (Walton, 1994). Proteases may make an important contribution by modifying the extensin network (and possibly determining the position of morphogenetic changes during plant development). This is particularly probable in the light of the model of the plant cell wall proposed by Talbott and Ray (1992) in which the plant cell wall is made up of cellulose microfibrils coated in hemicelluloses and other matrix polymers, with no direct microfibril cross-linking. Such a cell wall would inevitably rely heavily on many different noncovalent interactions between wall polymers. This model is an alternative to that proposed by McCann and Roberts (1991) where cellulose microfibrils are bound together by hemicelluloses, providing a strong structural framework.

Xyloglucan transglycosidase (XET) has been proposed to have wall-loosening activity (Fry *et al.*, 1992). It catalyzes a reaction in which a xyloglucan backbone is cut and one of the resulting chains is added to another xyloglucan chain. There is however much doubt about the impact XET action has on wall extensibility and morphology

(Cosgrove, 1998). Hydroxyl radicals can also loosen cell walls by causing oxidative cleavage of polysaccharides, particularly under acidic conditions (Fry, 1998).

The action of wall-loosening agents could be enhanced by changes in wall pH and secretion of activators and inhibitors of wall enzymes (Cosgrove, 1999). The extent to which each of these factors contributes to wall relaxation is not fully understood, although it seems that expansins play a major role and XETs only make a minor contribution (Cosgrove, 1998). Genetic engineering holds much potential in the future as a powerful method of investigating the architecture of the plant cell wall and the mechanism of cell wall expansion (Reiter, 1994).

1.8 Objectives of the project

Part of the basis of the proposed research extends from previous research performed at the University of Durham, which has determined an array of partial amino acid sequences of proteins found in the extracellular matrix of *Arabidopsis thaliana*, carrot, French bean, tomato and tobacco suspension cultured cells (Robertson *et al.*, 1997). The strategy adopted to obtain amino acid sequence data included differential extraction of the primary cell walls and resolution of these protein extracts and proteins present in the culture medium by SDS-PAGE. The N-terminal sequence of one of the proteins found in the culture filtrate of *Arabidopsis* suspension cultured cells displayed 100% identity to part of the predicted amino acid sequence encoded by *ara12*, a previously described cDNA sequence from *Arabidopsis* (Ribeiro *et al.*, 1995). *ara12* appears to encode a subtilisin-like protease and the N-terminal sequence obtained corresponds to the putative N-terminus of the mature protease. Plant

subtilisin-like proteases are a relatively poorly understood group of enzymes and their exact functions are either ill-defined or remain totally undefined. This project focuses principally on the nature and biological activity of the Ara12 protease and homologues of this protease in *Arabidopsis thaliana*.

A number of different subtilisin-like proteases appear to be found in *Alnus glutinosa* (alder; Ribeiro *et al.*, 1995) and *Lycopersicon esculentum* (tomato; Meichtry *et al.*, 1999). It has been proposed to determine whether a similar situation exists in *Arabidopsis*. If, as is suspected, a number of different subtilisin-like proteases can be detected in *Arabidopsis*, it is further proposed to isolate one or several of these by RT-PCR or cDNA library screening and use the cDNA sequences obtained to investigate the tissue expression of the corresponding genes by Northern analysis.

One of the specific aims of this work has been to examine the precise location of the Ara12 protease in the plant by immunocytochemistry. It is proposed to raise antisera to recombinant proteins carrying Ara12 epitopes overexpressed in *E. coli*. Establishing the cellular and subcellular locations of the Ara12 protease may help in determining the function of the enzyme. This has not been investigated before in *Arabidopsis*, although the expression pattern of *ag12* (a homologue of *ara12* identified from *Alnus glutinosa*) has been looked at in *Alnus glutinosa* root nodules using ^{35}S-labelled antisense and sense RNAs (Ribeiro *et al.*, 1995).

A further aim of the project has been to purify the Ara12 protease in its native form from *Arabidopsis* cell suspension cultures. The filtrate of these cultures is known to be rich in the mature Ara12 protease (Robertson *et al.*, 1997). Initially it will be determined whether a proteolytic activity can be detected in the culture filtrate using chromogenically and fluorogenically labelled casein substrates. In this way the subsequent purification of Ara12 protease, using conventional chromatographic

methods, can be monitored. The antisera generated during the course of this work will also be useful in purifying the protease or in following the protein purification.

Purification of the Ara12 protein from *Arabidopsis* cells will enable the first study of a subtilisin-like protease isolated from this model plant. Having purified the protease to homogeneity it is intended to determine the optimum conditions for its activity, including pH, temperature and a requirement for cofactors. Many subtilases are activated in the presence of cofactors, especially calcium ions. The inhibition of the hydrolytic activity of Ara12 will be examined using known serine protease inhibitors. The substrate specificity of the protease will be investigated using purified native proteins, cell wall extracts and artificial peptide substrates. This may give some insight into potential endogenous *in vivo* substrates of this enzyme.

Chapter 2

Materials and methods

2.1. Materials

2.1.1 General reagents

Solvents and chemicals were obtained from BDH Chemicals Ltd., Fisons, Gibco-BRL, Pharmacia, Sigma Chemical Corp. or Bio-Rad. DTT was obtained from Melford Laboratories Ltd. All reagents were used and stored according to the manufacturer's recommendations, unless stated otherwise.

2.1.2 Radiochemicals

[α-^{35}S]dATP (12.5 mCi/ml; 1200 Ci/mmol) and [α-^{32}P]dATP/dCTP (10 mCi/ml; 4500 Ci/mmol) were obtained from ICN and NEN Research Products.

2.1.3 X-ray film

Kodak X-OMAT AR and Fuji RX medical X-ray film was used respectively to detect [^{35}S] and [^{32}P] emissions. Film was exposed to radioactive or luminescent sources in light-tight cassettes at –70°C. Autoradiographs were developed and fixed, using Kodak developer and fixer, manually according to the manufacturer's instructions.

2.1.4 Solutions and buffers

1x TE buffer

10 mM Tris.HCl (pH7.4, 7.6, 8.0)

1 mM Na_2EDTA (pH8.0)

10x TES buffer

100 mM Tris.HCl (pH8.0)

10 mM Na_2EDTA (pH8.0)

1 M NaCl

50x TAE

Tris base 242 g

Glacial acetic acid 57.1 ml

Na_2EDTA 18.6 g

Made up to 1 litre with distilled water.

TTB buffer

Tris base 3.03 g

glycine 14.4 g

methanol 200 ml

Made up to 1 litre with distilled water.

TBS:Tween 20

10 mM Tris.HCl (pH8.0)

0.15 M NaCl

0.5 ml Tween 20 (polyoxyethylene sorbitan monolaurate)

Made up to 1 litre with distilled water.

Solution I

50 mM glucose

25 mM Tris.HCl (pH8.0)

10 mM Na_2EDTA (pH8.0)

Solution II

0.2 M NaOH

1% (w/v) SDS

This solution was freshly prepared before use by mixing equal volumes of 0.4 M NaOH and 2% (w/v) SDS stock solutions.

Solution III

60 ml 5 M potassium acetate

11.5 ml glacial acetic acid

28.5 ml distilled water

Plant cell lysis buffer

50 mM Tris.HCl (pH7.6)

100 mM NaCl

50 mM Na_2EDTA

0.5% (w/v) SDS

10 mM 2-mercaptoethanol

Mix and adjust to pH7.6.

DNA loading dye

15% (w/v) Ficoll 400

0.25% (w/v) bromophenol blue

NaOH was added dropwise until the pH change caused the solution to turn deep blue.

Ethidium bromide

EtBr was kept as a 10 mg/ml stock solution in distilled water and was kept in the dark.

5x MOPS buffer

0.2 M MOPS

50 mM sodium acetate

5 mM Na_2EDTA

Adjust to pH7.0 using NaOH and add 0.1% DEPC. Leave overnight and autoclave.

Depurination solution

0.25 M HCl (corresponds to 21.6 ml of 11.6 M stock HCl diluted to 1 litre with MilliQ water).

Denaturing solution

1.5 M NaCl

0.5 M NaOH

Neutralizing solution

1.5 M NaCl

0.5 M Tris.HCl (pH7.4)

20x SSC

3 M NaCl

0.3 M sodium citrate, pH7.0

20x SSPE

3 M NaCl

0.2 M $NaH_2PO_4.H_2O$

20 mM EDTA

Adjust to pH7.4 and make up to 1 litre with distilled water.

50x Denhardt's solution

1% (w/v) Ficoll 400

1% (w/v) polyvinylpyrrolidone

1% (w/v) BSA

Made up with distilled water, filtered and stored in aliquots at -20°C.

dNTP stock solution

25 mM dATP, 25 mM dCTP, 25 mM dGTP, 25 mM dTTP in distilled water

Kept frozen in aliquots at -20°C.

Lowry solution A

2% (w/v) Na_2CO_3

0.1 M NaOH

Lowry solution B

0.5% (w/v) $CuSO_4$

1% (w/v) sodium citrate

SDS-PAGE running buffer

25 mM Tris base

192 mM glycine

0.1% (w/v) SDS

5x SDS-PAGE sample buffer

10% (w/v) SDS

5% (w/v) DTT

0.05% (w/v) bromophenol blue

0.3 M Tris.HCl, pH6.8

50% (v/v) glycerol

Coomassie staining solutions

Coomassie I

10% glacial acetic acid

25% propan-2-ol

0.25% Brilliant Blue Coomassie R-250

Coomassie II

10% glacial acetic acid

10% propan-2-ol

0.03% Brilliant Blue Coomassie R-250

Coomassie III

10% glacial acetic acid

0.03% Brilliant Blue Coomassie R-250

Destain

10% glacial acetic acid

1% glycerol

Silver staining solutions

Fixing solution

40% (v/v) ethanol

10% (v/v) acetic acid

Incubation solution

75 ml ethanol

17.0 g sodium acetate

1.3 ml glutaraldehyde (25% w/v)

0.5 g sodium thiosulphate

Made up to 250 ml with deionized water.

Silver solution

0.25 g silver nitrate

50 µl formaldehyde (40% v/v)

Made up to 250 ml with deionized water.

Developing solution

6.25 g sodium carbonate

25 µl formaldehyde (40% v/v)

Made up to 250 ml with deionized water.

Ponceau S

0.2% (w/v) Ponceau S stain (sodium salt of 3-hydroxy-4-[2-sulpho-4-(4-sulphophenylazo) phenylazo]-2,7-naphthalenedisulphonic acid) in 1% (v/v) acetic acid.

2.1.5 Chromatography columns and resins

Q Sepharose FF (fast flow) and phenyl Sepharose FF, arginine Sepharose 4B, benzamidine Sepharose 4B, as well as Mono-Q, Mono-S, phenyl Superose, C8 RP-HPLC and PD-10 pre-packed columns were obtained from Amersham Pharmacia LKB Biotechnology. Casein agarose was purchased from Sigma.

2.1.6 Enzymes

Restriction enzymes and other DNA modifying enzymes were purchased from Boehringer Mannheim, New England Biolabs (NEB), Northumbria Biologicals Ltd. (nbl) or Promega Ltd. The Rediprime priming kit was obtained from Amersham. The Random Primed DNA Labeling Kit was supplied by Boehringer Mannheim. *Taq* DNA polymerase was obtained from Bioline.

Subtilisin Carlsberg was purchased from Sigma and Factor Xa protease was purchased

from New England Biolabs.

Purchased enzymes were used with the manufacturer's buffer unless otherwise stated (see Table 2.1 for restriction enzyme buffers). Enzymes and their buffers were stored, in aliquots if appropriate, at -20°C or 4°C.

2.1.7 Nucleic acids

DNA molecular weight markers were resolved on agarose gels alongside linear DNA samples to assess the approximate size of DNA fragments. Bacteriophage lambda DNA digested with the restriction enzymes *Hin*dIII and/or *Eco*RI and ϕX174 DNA digested with *Hae* III was obtained from Boehringer Mannheim and nbl. Herring sperm DNA (10 mg/ml) was sheared by passing it through a 19 G needle 12 times and was used in prehybridization solutions at a concentration of 100 μg/ml.

Oligonucleotides used in this work were synthesized by MWG and Cruachem and are listed in Table 2.2.

10x buffer	Composition
A	33 mM Tris.acetate (pH7.5), 10 mM magnesium acetate, 66 mM potassium acetate, 0.5 mM DTT.
B	10 mM Tris.HCl (pH8.0), 5 mM $MgCl_2$, 100 mM NaCl, 1 mM 2-mercaptoethanol.
H	50 mM Tris.HCl (pH7.5), 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTT.
L	50 mM Tris.HCl (pH7.5), 10 mM $MgCl_2$, 1 mM DTT.
M	50 mM Tris.HCl (pH7.5), 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT.

Table 2.1 Restriction enzyme buffers.

Ref. no.	DNA sequence
SUB278	5'- GGAATTCCATATGCGTTACGAGCTACACACC -3'
SUB2239	5'- AAGGAAAAAAGCGGCCGCTGTCCAGCTAATCGCCACGG -3'
MAL239	5'- ATGACTCAACCTGGTGTTATCTCC -3'
pMALFOR	5'- GGTCGTCAGACTGTCGATGAAGCC -3'
SUB1597	5'- ATGTCTTGCCCTCACGTTAGTGG -3'
MAL2247	5'- CTAGTCTAGATACGACTATGTCCAGCTAATCGC -3'
Slpa5' (SUB5F)	5'- CAACACACCATCTCAATCATGTC -3'
Slpa3' (AS2244)	5'- GGGGTACCCGACTATGTCCAGCTAATCGG -3'
Homb5'	5'- AGAATTCCAATGGCTTCCTCCACCATC -3'
Homb3'	5'- CGAATTCGAATCACAACGTATCCATTTGG -3'
Homc5'	5'- AGAATTCCTCCAATGGAACCCAAACCTTTC -3'
Homc3'	5'- CGAATTCGCTCAGTTAGTCTTCAAGGTTAC -3'
Homd5'	5'- TGAATTCAGATATGCTTCAGCCGAAGATG -3'
Homd3'	5'- GGAATTCATTAATTTCAATCCGAAGTAGGAC -3'
Homg5'	5'- CGAATTCCCAAAATGGATATCGGGTTGAG -3'
Homg3'	5'- AGAATTCGATTTGTTCATCGCCGATGTCC -3'
Homh5'	5'- TGAATTCGGATGGCTATAGCTTTTCATACC -3'
Homh3'	5'- GGAATTCACATGCAAGCTTTTGG -3'
Homj5'	5'- TGAATTCAGCTATGGCGTTAGAAGCAAAG -3'
Homj3'	5'- GGAATTCGGTTTTAACTGACAACGGTATTC -3'
Homk5'	5'- GGAATTCTACTCTAGCTCCAACAATGGC -3'
Homk3'	5'- TGAATTCGTTCTACAGAGGAGGCAACC -3'
STOM1F	5'-GGAATTC-AC/TAAC/TCAA/GTGC/TGGIGTITGGGC-3'
STOM2R	5'-GGAATTC-GCIGTA/GTGIGTICCA/GTGA/C/G/TCC-3'
STOM2F	5'-GGAATTC-CAC/TACIA/TC/GIGAC/TTTC/TC/TTIAAA/GC/TT-3'

Table 2.2 Oligonucleotides used in this work.

2.1.8 Bacterial strains

The strains used in this work are listed with their genotypes in Table 2.3. Cultures were stored at 4°C on agar plates wrapped in Parafilm, or, for long-term storage, at -70°C in a final concentration of 15% (v/v) glycerol.

***Escherichia coli* strain**	**Genotype**
DH5α	ϕ80d*lacZ*ΔM15, *recA1*, *endA1*, *gyr A96*, *thi-1*, *hsdR17* (r_K^-,m_K^+), *supE44*, *relA1*, *deoR*, Δ(*lacZYA-argF*)) U169
BL21 (DE3)	F$^-$, *ompT*, *hsdS*$_B$, (r_B^-, m_B^-), *dcm*, *gal*, λ(DE3)
XL1Blue	Δ(*mcrA183*, Δ(*mcrCB-hsdSMR-mrr)173*, *endA1*, *supE44*, *thi-1*, *recA1*, *gyrA96*, *relA1*, *lac[F' proAB*, *lacI*^{q}ZΔM15, Tn109tet^r)]
JM103Y	*Dlac-pro*, *thi*, *strA*, *supE*, *endA*, *sbcB15*, *hsdR4*, *F'[traD36*, *proAB+*, *lacI*^{q}, *lacZDM15]*

Table 2.3 Bacterial strains used in this work.

2.1.9 Plants and plant cell suspension cultures

The plant used in this work was *Arabidopsis thaliana* Ecotype Columbia. Previously described *Arabidopsis thaliana* and tomato (*Lycopersicon esculentum*) cell suspension cultures were used in this work (Robertson *et al.*, 1997).

2.1.10 Plant tissue for immunocytochemistry

Arabidopsis tissue embedded in resin was kindly supplied by Dr. Jackie Spence. Freshly excised tissue was fixed for 12 h at room temperature in 3% (w/v) paraformaldehyde, 1.25% (w/v) glutaraldehyde, 50 mM phosphate buffer pH7.0. Tissues were sequentially dehydrated for 1h in a graded series of different ethanol solutions: 12.5%, 25%, 50%, 75% and 95% ethanol. Tissues were then embedded in London Resin LR White, an aromatic acrylic resin (Hall and Hawes, 1991). A heat polymerization step was performed at 50°C.

2.1.11 Media

Media solids were obtained from Difco. Media was prepared as described by Sambrook *et al.* (1989) and sterilized by autoclaving for 20 min at 15 psi. If required, supplements were added to the media from separate sterile stock solutions. Antibiotics were added to liquid and agar media after autoclaving to select for a plasmid carrying a gene conveying resistance to that antibiotic. The following antibiotics were filter sterilized through a 0.2 or 0.4 μm filter and used in media at the quoted concentrations:

100 mg/ml ampicillin (1000x stock; final concentration 100 mg/ml)

12.5 mg/ml kanamycin (500x stock; final concentration 25 mg/ml).

2.1.11.1 Media for bacteria

LB (Luria-Bertani) medium

1% (w/v) Bacto-tryptone

0.5% (w/v) Yeast extract

0.25% (w/v) NaCl

L-plate agar

1% (w/v) Bacto-tryptone

0.5% (w/v) Yeast extract

0.25% (w/v) NaCl

1.8% (w/v) Agar

2xYT medium

1.6% (w/v) Bacto-tryptone

1% (w/v) Yeast extract

0.5% (w/v) NaCl

Terrific broth

1.2% (w/v) Bacto-tryptone

2.4% (w/v) Yeast extract

0.004% (v/v) Glycerol

17 mM KH_2PO_4, 72 mM K_2HPO_4

SOC Medium

SOC medium was made by mixing equal volumes of the sterile stock solutions 2xSOC salts and 2xTYEG.

2xSOC salts

20 mM NaCl

5 mM KCl

20 mM $MgCl_2$

20 mM $MgSO_4$

2xTYEG

4% (w/v) Bacto-tryptone

1% (w/v) Yeast extract

This solution was autoclaved and then sterile glucose was added to a final concentration of 40 mM.

2.1.11.2 Media for plant cell suspension cultures

Plant cells were maintained in the appropriate plant cell suspension culture media under a 16 h photoperiod at 24°C at 130 rpm. Suspension cultures were subcultured once every 7 days by aseptically transferring 10 ml of culture to 90 ml of fresh media in 250 ml Erlenmeyer flasks. Plant callus was maintained on plant callus agar plates under a 16 h photoperiod at 24°C and was subcultured to fresh agar plates every 1-2 weeks.

Arabidopsis cell suspension culture media

3% (w/v) sucrose

0.44% MSMO (Murashige and Skoog basal salts with minimal organics)

0.5 mg/l NAA (naphthalene acetic acid), 0.05 mg/l kinetin

Adjust to pH5.7 with 1 M NaOH.

Tomato cell suspension culture media

2% (w/v) glucose

0.43% MS (Murashige and Skoog basal salts)

2 mg/l glycine

100 mg/l myo-inositol, 0.5 mg/l nicotinic acid, 0.5 mg/l pyridoxine/HCl, 0.1 mg/l thiamine

10 mg/l indole acetic acid, 0.5 mg/l kinetin

Adjust to pH5.7 with 1 M NaOH.

Plant callus agar

This was made by adding 0.6% Difco Bactoagar to the plant cell suspension culture media described above.

2.2 Methods

2.2.1 DNA and RNA related methods

DNA was extracted or manipulated either in clean sterile plasticware or in glassware and plasticware which had been soaked in 0.1 M HCl overnight.

2.2.1.1 Equilibration of phenol

Redistilled phenol (BDH) was extracted once with each of the following solutions: 1 M Tris.HCl (pH8.0), 0.1 M Tris.HCl (pH8.0) and 0.2 M 2-mercaptoethanol. The phenol was extracted twice with an equal volume of 1xTE buffer (pH8.0), mixed with 0.25 volumes of 1xTE buffer (pH8.0) and stored at 4°C in the dark. Pre-equilibrated phenol was also obtained from Appligene.

2.2.1.2 Phenol extraction of aqueous DNA solutions

Tris-saturated phenol was used to denature and subsequently remove protein impurities from aqueous DNA solutions. An equal volume of phenol was added to a dilute DNA solution, vortex mixed for 1 min and centrifuged for 5 min in a microfuge. The upper aqueous phase was transferred to a clean tube, taking care not to dislodge the protein interface between the two layers. This extraction process was repeated with equal volumes of phenol:chloroform:isoamyl alcohol 25:24:1 (v/v) and lastly with chloroform: isoamyl alcohol 24:1 (v/v). Chloroform helps to remove lipids from the aqueous layer and isoamyl alcohol is used as an anti-foaming agent.

2.2.1.3 DNA precipitation

To precipitate DNA from an aqueous solution, 0.1 volumes 3 M sodium acetate (pH6.5) and 2-3 volumes of absolute ethanol were added. The solutions were mixed and kept either at -70°C for 20 min or at -20°C for 2 h. The resultant precipitate was spun down at 13,000 x g for 10 min and the supernatant was discarded. The pellet was washed with 70% ethanol and centrifuged as before. The supernatant was removed by aspiration and by aid of a tissue. The DNA was resuspended in 1xTE buffer (pH7.6) or sterile distilled water. The precipitation procedure was also performed using propan-2-ol rather than ethanol.

2.2.1.4 Alkaline lysis mini-preparation of plasmid DNA from *E. coli*

LB medium (2 ml), containing appropriate selective antibiotics, which had been inoculated with *E. coli* carrying the plasmid to be isolated, was grown overnight at 37°C on a rotary shaker. A 1.5 ml aliquot of the overnight culture was transferred to an Eppendorf tube and centrifuged for 5 min in a microfuge. The supernatant was discarded and the pellet was resuspended in 100 µl ice-cold solution I by vortexing (Birnboim and Doly, 1979). After this, 200 µl solution II was added. The tube was inverted several times and incubated on ice for 5 min, before adding 150 µl solution III. The mixture was vortexed briefly, kept on ice for a further 5 min and centrifuged for 5 min. This process separated a flocculent precipitate from the supernatant and the latter was removed to a clean tube. An equal volume (450 µl) of phenol:chloroform:isoamyl alcohol 25:24:1 (v/v) was added, vortexed for 1 min and spun for 5 min. The upper aqueous phase which formed was transferred to a tube containing 45 µl 3 M sodium acetate pH6.5 and 1 ml absolute ethanol. This tube was

kept at -70°C for 30 min and was then centrifuged for 10 min in a microfuge. The pellet was washed in 0.5 ml 70% ethanol and spun in the microfuge for 10 min. The bulk of the ethanol was removed by aspiration using a Gilson micropipette, whilst the remaining ethanol was removed using a tissue. The final pellet was dried under vacuum and resuspended in 30 µl 1xTE buffer (pH7.6) or sterile distilled water.

Plasmid DNA was also prepared from *E. coli* cultures using the Hybaid Recovery Plasmid Prep kit for small scale preparations of high purity DNA, particularly for nucleotide sequencing and the QIAGEN Plasmid Midi kit for larger scale preparations, useful for producing cloning quality DNA. The manufacturer's protocols were adhered to throughout.

2.2.1.5 Rapid preparation of plant DNA

Using a pestle and mortar, 50-100 mg of fresh leaf tissue was ground up in liquid nitrogen. This material was suspended in 1.5 ml plant cell lysis buffer and left at room temperature for 15 min. The lysate was split into two equal aliquots, using a 1 ml pipette tip with a cut end. To both aliquots 300 µl phenol/Tris.HCl pH7.5 was added and vortexed for 1 min. After leaving to stand at room temperature for 2 min, 300 µl chloroform was added. The mixture was vortexed briefly and spun in a microcentrifuge for 2 min. The upper aqueous phase underwent another phenol/chloroform extraction as before. In order to remove any remaining phenol an equal volume of chloroform was mixed with the supernatant and centrifuged for 2 min. An equal volume of isopropanol was added and incubated at room temperature for 5min. The DNA precipitate was spun down for 5 min and the pellet was washed with 1 ml 70% ethanol. The DNA was resuspended in 1xTES

buffer containing 50 µg/µl RNase and was incubated for 30 min at 37°C. At this stage the RNase was removed by one phenol/chloroform extraction followed by a chloroform extraction. The DNA was precipitated and the pellet was resuspended in 100 µl TE buffer (pH7.6). Total plant DNA isolated by this method was suitable for restriction and hybridization studies (Junghans and Metzlaff, 1990).

2.2.1.6 Total RNA preparation from plant tissue

Total RNA was prepared from different plant tissues using Trizol reagent from Gibco BRL according to the manufacturer's instructions. Total RNA preparations were stored at -80°C.

2.2.1.7 Estimation of DNA and RNA concentrations

DNA and RNA concentrations of aqueous solutions were determined by measuring the absorbance values at 260 nm and 280 nm (A_{260} and A_{280}) using a spectrophotometer against an appropriate blank cuvette. The A_{260} value gave an estimate of the concentration of the nucleic acid. An A_{260} value of 1.0 corresponds to:

33 mg oligonucleotide / ml

40 mg RNA /ml

50 mg double stranded DNA (plasmid/genomic) / ml

The A_{260}:A_{280} ratio provided an indication of the purity of nucleic acids in aqueous solution. Any phenol or protein present in the nucleic acid solution compromises the purity and reduces these ratios from 1.91, for a pure DNA solution and 2.03, for a pure RNA

solution. The concentration of linear DNA fragments was also assessed by comparison with DNA standards (usually λ DNA digested with *Hin*dIII) of known concentration on agarose gels.

2.2.1.8 Heat treatment of RNase

RNase A (10 mg/ml) from bovine pancreas was stored at -20°C after boiling for 10 min to destroy any possible DNase activity.

2.2.1.9 DNA digestion using restriction endonucleases

The recommendations of the manufacturer were observed for each enzyme used. Different restriction endonucleases may require varying conditions for maximum efficiency. Optimum pH, salt content, 2-mercaptoethanol content, the temperature of incubation and other factors varied between the enzymes. A range of 10x reaction buffers were used (Table 2.1), which conferred known efficiencies to the restriction enzymes.

One unit (1 U) of enzyme is defined as the amount of enzyme required to digest 1 μg λ DNA completely under optimum conditions in 1 h. Reaction mixtures were normally incubated with appropriate amounts of restriction enzymes at 37°C for 1-2 h. Double digests (using two different restriction enzymes) were performed concurrently if the buffer requirements proved to be compatible. If these requirements were incompatible, then the digests were carried out sequentially, ethanol precipitating the DNA between digestions.

2.2.1.10 Agarose gel electrophoresis of DNA

An approximate guide to the appropriate percentage agarose gels (prepared with 1x TAE) to use to resolve DNA fragments was obtained from Sambrook *et al.* (1989) as shown in Table 2.4. To ensure that the agarose was completely dissolved in 1x TAE it was heated in a microwave. The intercalating dye ethidium bromide was added to a final concentration of 0.5 mg/ml. The running buffer used was 1x TAE. Gels were run at 120 V respectively for 30 min-1 h unless otherwise stated. The DNA samples were supplemented with 0.2 volumes DNA loading dye before loading onto the gel and applying the current. Resolved DNA fragments were visualized using a UV transilluminator.

2.2.1.11 Purification of DNA fragments from agarose

An agarose gel slice was excised and most of the excess agarose was removed. The gel slice was placed inside 0.5 ml Eppendorf tubes, which had been holed with a 19 G needle and plugged with a small amount of glass wool. These small tubes were placed inside 1.5 ml Eppendorf tubes and spun in a microfuge for 1 min. The eluate was pooled, including any pieces of agarose present in the large Eppendorfs. The volume of each pool was made up to 0.5 ml with distilled water and extracted once with phenol/chloroform. The supernatant DNA was precipitated by adding 50 µl 3 M sodium acetate pH6.5 and 1 ml absolute ethanol. The tubes were stored at -70°C for 1 h and then spun for 15 min. The pellet was washed with 500 µl 70% ethanol and spun for 10 min. After discarding the ethanol the tubes were given a 10 sec spin. The last traces of ethanol were removed with a micropipette and the pellet was resuspended in 20 µl sterile distilled water. To check the

Percentage agarose (w/v)	Efficient range of separation of linear DNA fragments (kb)
0.3	5 – 60
0.6	1 – 20
0.7	0.8 – 10
0.9	0.5 - 7
1.2	0.4 - 6
1.5	0.2 - 3
2.0	0.1 - 2

Table 2.4 Practical range of DNA fragment separation varying with agarose concentration. (Table taken from Sambrook *et al.*, 1989).

recovery of DNA 2-10 µl of this solution was run on an agarose gel.

QIAEX spin columns were also used for efficient recovery of resolved DNA fragments from agarose gel. In this method solubilized DNA binds to a silica based resin in a

chaotropic solution at neutral pH. Columns were supplied by QIAGEN and the instructions for gel fragment QIA-prep purification were adhered to, except that EDTA was omitted from the wash buffers.

2.2.1.12 Labelling DNA by random priming

This procedure was developed by Feinberg and Vogelstein (1984). DNA was radiolabelled using a Boehringer Mannheim random primed DNA labelling kit. If DNA contained in an agarose gel slice was to be labelled, the gel slice was boiled for 7 min and then incubated at 37°C for 15 min. If the DNA to be labelled was in aqueous solution, then the DNA was diluted with distilled water to approximately 2.5 ng DNA / µl. The molten gel slice (11 µl) or the aqueous DNA solution (11 µl) was put in a fresh Eppendorf tube in which the labelling reaction was performed. In addition 1 µl distilled water, 2 µl reaction mix (random hexanucleotide primer) and 1 µl each of the 0.5 mM 2'-deoxynucleoside-5'-triphosphate (dNTP) solutions dCTP, dGTP and dTTP were added. Lastly, 2 µl [α-^{32}P]dATP and 1 ml Klenow enzyme were mixed in and incubated at 37°C for 1 h on a heat block using protective lead shielding. 80 µl 1xTE buffer (pH7.6) was mixed with the solution, which was then centrifuged through a Sephadex G-50 column for 1 min at 1600 x g. This was done so that most of the unincorporated nucleotides would be retained in the column. The oligonucleotide probe was boiled for 10 min and snap cooled on ice prior to use in hybridization experiments.

The Rediprime priming kit (Amersham) was also routinely used to label DNA probes. Random octomers annealed to the denatured probe DNA and a radiolabelled second strand

was synthesized incorporating [^{32}P]-labelled 2'-deoxycytidine 5'-triphosphate ([α-^{32}P]dCTP).

2.2.1.13 Sephadex G-50 column chromatography

This method of filtration was used to separate DNA species of differing molecular weights. For example, small sticky ended fragments of DNA released during a restriction digest can be removed by this technique. A sterile 1 ml disposable syringe was plugged with glass wool and filled to the top with a sterile suspension of 3 g Sephadex G-50 (or G-25) beads (Pharmacia) suspended in 100 µl 1xTE buffer (pH7.6). The syringe was placed in a Falcon tube and spun at 1600 x g for 1 min. Sephadex G-50 was added as before and the column was spun for 1 min to compact it. 100 µl sterile 1xTE buffer was spun through the column at 1600 x g for 1 min. The DNA sample was made up to a volume of 100 µl with 1xTE and spun through the column.

2.2.1.14 Southern blotting

This method was used to transfer and immobilize DNA onto a solid support for subsequent analysis (Southern, 1975). Genomic DNA (5-10 mg) was digested with restriction enzymes and the DNA fragments were resolved on a 0.7% agarose gel. The gel was taken out of the gel tank and washed for 30 min on a Denley shaker in 0.25 M HCl to cause DNA depurination. This partially cleaves large DNA fragments *in situ*, assisting in an efficient transfer. The gel was then washed with rocking in denaturing solution for 45-60 min and then in neutralizing solution for 1 h.

About 100 ml 20xSSC was poured into a plastic container and a glass plate, large enough to accommodate the gel was put on the container. A long wick of Whatman 3MM paper was moistened with the 20xSSC and laid over the glass plate, dipping into the solution on either side. The bottom (anode end) right corner of the gel (as viewed from above as the gel was running) was cut off, the gel was turned upside down and placed on top of the wick. A nylon filter of the same size as the gel (and with a corresponding corner cut off) was wetted with 20xSSC and placed over the gel. Any air bubbles between the layers were removed, making sure that gloves were worn. A layer of wet 3MM paper was placed onto the membrane, followed by a stack of absorbant paper towels and a heavy weight. Parafilm was positioned around the gel edges to prevent a "short circuiting" effect of the flow of solution from the container through the wick, gel and membrane and into the paper towels. DNA transfer to the hybridization membrane usually occurred overnight.

After blotting, the membrane was air dried. Nylon membranes were UV crosslinked (using 120 mJ/cm^2) in a Stratalinker (Stratagene) for 20 sec. After fixing the DNA the membranes were ready for hybridization and were stored between sheets of Whatman 3MM paper in sealed plastic bags at 4°C. After hybridization membranes were not allowed to dry out.

2.2.1.15 Electrophoresis of RNA and Northern blotting

Total RNA was resolved on 1% agarose gels containing formaldehyde. 1.5 g agarose was dissolved in 108 mls of hot sterile distilled water, which had been treated with 0.1% diethyl pyrocarbonate (DEPC). This was cooled to 70°C before adding 12 mls formaldehyde and

30 mls 5x MOPS (3-(*N*-morpholino) propanesulphonic acid) which had been heated to 60°C. The gel was poured in a fume hood. Gels were pre-run for 30 min at 80-100 V in 1x MOPS prior to loading.

An equal volume of freshly prepared denaturing buffer (containing 200 µl 5x MOPS, 120 µl formaldehyde and 100 µl deionised formamide), was added to the RNA samples. 1 µl 5 mg/ml EtBr was also sometimes added to each sample. Samples were heated to 65°C for 10 min, cooled on ice and 0.1 volumes of loading buffer was added before loading.

After electrophoresis the formaldehyde gel was rinsed several times. RNA was transferred and fixed to Hybond-N nylon filters (Amersham) as described for the Southern blotting procedure, omitting the depurination, denaturation and renaturation steps.

2.2.1.16 Colony blotting

A nylon hybridization membrane of the same size as the agar plate was placed on to the agar surface for 1-2 min. Colonies were lifted directly onto the filters (Grunstein and Hogness, 1975). Three asymmetric holes were made in the membrane with a sterile needle, so that the filter orientation could be ascertained. The discs were removed with sterile forceps and placed colony side up in turn onto Whatman 3MM paper saturated with 10% SDS for 3 min, denaturing solution for 5 min, neutralizing solution for 5 min and 2xSSC for 5 min. Filters were dried briefly between each step. The filters were air dried for 30 min, sandwiched between two dry sheets of 3MM paper and glass plates and placed in a UV cross-linker for 20 sec. Fixed membranes should be prewashed prior to hybridization to remove colony debris, reducing background noise.

2.2.1.17 Hybridization of nucleic acid probes

Hybridization membranes were incubated with prewarmed SSC or SSPE based prehybridization solution containing Denhardt's solution, SDS and denatured herring sperm DNA (see Sambrook *et al.*, 1989 for details) in rotating bottles for 2 h at 65°C for DNA samples and 42°C for RNA samples. The hybridization steps were performed in a Hybaid Hybridization Oven with a rotisserie. The radiolabelled probe was denatured (by boiling for 10min) and chilled on ice. The prehybridization solution was poured off and prewarmed SSC or SSPE based hybridization solution containing Denhardt's solution, SDS and denatured herring sperm DNA (see Sambrook *et al.*, 1989 for details) containing the probe was added to the hybridization bottle. The filter was incubated on a rotisserie overnight at varying temperatures according to the stringency required (high stringency incubations were performed at 65°C and 55°C for DNA and RNA samples, respectively).

The solution was discarded via a radioactive sink or kept for reuse. The filter was usually washed at least twice for 20 min in 0.1% (w/v) SDS solutions containing decreasing amounts of SSPE or SSC. The stringency of the washes depends on their duration and the salt concentration and temperature of the wash solution. Excess liquid was blotted with filter paper and the membranes were wrapped in cling film and exposed to X-ray film, between two intensifying screens, at -80°C overnight. The membrane was not allowed to dry out and could then be reprobed. Filters were stripped of the probe by boiling in 0.1xSSC, 0.1% SDS for 10 min.

2.2.1.18 Removal of 5' phosphate groups

In preparation for ligation reactions, linear vector DNA was treated with phosphatase to remove 5' phosphate groups. This favours intermolecular ligation events over intramolecular ones, thereby promoting the success of a ligation reaction. One unit (1 U) of alkaline phosphatase (from calf intestine or shrimp) and 0.1 volumes of the 10x phosphatase buffer supplied was added per 1 μg of linear double stranded DNA. This was incubated at 37°C for 10 min. The enzyme activity was then destroyed by incubating the mixture at 70°C for 10 min. Occasionally the DNA was then phenol extracted and precipitated using 0.5 volumes 7.5 M ammonium acetate and 3 volumes absolute ethanol. The DNA was washed in 70% ethanol and resuspended in sterile distilled water to concentrations of 0.1-0.5 μg/μl for use in ligation reactions.

2.2.1.19 DNA ligation

Ligation reactions were carried out in a total volume of 10 or 20 μl with a total of 100-500 ng DNA. One unit of bacteriophage T4 DNA ligase and 0.1 volumes 10x ligation buffer were added to the DNA. If necessary the reaction volume was made up to the required amount with sterile distilled water. Ligation reactions were placed in a 14°C incubator for 2 h or overnight.

Condensing agents, like 15% PEG (polyethylene glycol) 8000 or 1-1.5 μM hexamminecobalt chloride can increase blunt-ended ligations 10 to 100-fold and 50-fold, respectively.

2.2.1.20 DNA amplification by PCR

Reaction mixtures were assembled in 0.5 ml Eppendorf tubes placed on ice. Solutions were added in the following order: MilliQ water, magnesium chloride, 10x*Taq* buffer, template DNA, primers, dNTPs and *Taq* DNA polymerase. Unless otherwise stated the final concentrations used in each amplification reaction were as follows: 1.5 mM $MgCl_2$, 1x*Taq* polymerase buffer (10 mM Tris-HCl pH9.0, 50 mM KCl, 0.1% Triton X-100, 0.01% gelatine), 0.2-0.4 μM forward and reverse primers, 0.25 mM of each deoxynucleoside triphosphate, 10 U/ml Taq polymerase (Bioline) and an appropriate of template (e.g. between 0.05-2 μg genomic DNA per 50 μl reaction) made up to 50 or 100 μl with sterile MilliQ water. Vent polymerase (New England Biolabs) with its superior proof-reading qualities was sometimes used instead of *Taq* DNA polymerase. In this case a Vent buffer was used and an extension temperature of 74°C was used.

The tubes with contents assembled were mixed, spun down briefly and one drop of mineral oil was added to each tube to reduce evaporation during the cycling. The tubes were incubated in a Techne thermocycler, according to the following cycling conditions, unless otherwise stated.

Stage 1 94°C for 5 min (initial denaturation)

Stage 2 94°C for 1 min (denaturation)

48-62°C for 1 min (annealing)

72°C for 1.5 min (extension)

30-40 cycles.

Stage 3 72°C for 10 min (final extension)

For DNA amplification from RNA templates the transcripts were first reverse transcribed. A 1st Strand cDNA Synthesis kit from Boehringer Mannheim or reverse transcriptase from Promega Ltd. was used for generating templates for RT-PCR from total RNA using oligo-dT primers.

2.2.2 Microbiological methods

2.2.2.1 Preparation of competent *E. coli* cells

LB medium (250 ml) was inoculated with 1/100th volume of a fresh *E. coli* overnight culture, which had been grown under antibiotic selection if possible. The cells were grown without selection for about 5 h at 37°C, with vigorous shaking, to an A_{600} of 0.5-1.0. The culture was chilled on ice for 30 min and centrifuged for 15 min at 4400 rpm in a J2-21 Beckman centrifuge using a JA-10 rotor which had been cooled to 4°C. As much of the supernatant as possible was removed and the cells were resuspended in 0.5 volumes of ice cold sterile 50 mM $CaCl_2$. The centrifugation step was repeated, and the pellet was resuspended in 0.1 volumes of ice cold sterile 50 mM $CaCl_2$, 20% (v/v) glycerol. This suspension was dispensed into 50-100 µl aliquots. If stored at -80°C these competent *E. coli* cells remain competent for transformation for at least 6 months.

2.2.2.2 Transformation of *E. coli* cells

If frozen, the competent *E. coli* cells required for this protocol were thawed on ice. A 50-100 µl aliquot of competent *E. coli* cells was mixed inside a pre-cooled Falcon tube with approximately 100-200 ng of DNA on ice for 30 min to 1 h. Inclusion of 25 mM 2-mercaptoethanol can increase transformation efficiencies by up to 3-fold. Cells were then heat shocked at 42°C for 30-40 sec and placed on ice. 1 ml SOC medium or 2xYT was added into the tube, which was then incubated at 37°C with shaking for 1 h.

The cells were centrifuged at 2000 rpm for 5 min in a bench centrifuge and resuspended in 100 µl LB medium. This suspension was spread onto selective LB agar plates and incubated at 37°C overnight. The appearance of *E. coli* colonies indicated the growth of transformed *E. coli* cells on the selective medium. Where possible blue/white selection of transformants was used using X-gal and IPTG (Sambrook *et al.*, 1989).

2.2.3 Protein related methods

2.2.3.1 Protein extraction

Unless otherwise stated the following method was used for extraction of *Arabidopsis* proteins. Using a pestle and mortar 100-500 mg of *Arabidopsis* tissue or whole plants were ground up in liquid nitrogen. This material was then homogenized in an ice cold solution of 20 mM HEPES, pH7.6, 100 mM potassium acetate using a glass teflon homogenizer. The homogenate was then filtered through three layers of Miracloth and used in Western blot analysis or in protein purification steps.

2.2.3.2 Plant cell wall (protein) extraction

Cell walls and extracellular protein extracts were generated using *Arabidopsis* suspension cultured cells 5 days after subculture. Cell wall material was prepared by passing these cells through a French press and layering the homogenised cells onto 10% (v/v) glycerol. Cell walls sedimented to the bottom of the glycerol solution and were extracted with ice cold 0.2 M $CaCl_2$ for 30 min. Extracellular protein extracts were generated according to the method of Robertson *et al.* (1997). Suspension cultured plant cells were harvested 5 days after subculture by filtration on Miracloth. Whole cells or isolated cell walls were washed 3 times in ice cold distilled water (3 ml/g fresh weight). Cells or cell wall material was stirred in 3 volumes of 0.2 M $CaCl_2$ for 30 min at 4°C, collected by filtration on Miracloth and washed 3 more times with distilled water as before. Subsequent extractions were carried out sequentially on the same cells for 30 min each, with 50 mM CDTA in 50 mM sodium acetate pH6.5, followed by 2 mM DTT, 1 M NaCl and finally 0.2 M borate pH7.5 at room temperature. Between extractions the cells or cell walls were washed on the filter 3 times, each with 3 volumes of distilled water. Extracts and the culture media were re-filtered through GF/A paper before being dialyzed against a 10-fold excess of distilled water with three changes.

2.2.3.3 Lowry and Bradford protein assays

Samples were assayed for protein by the Lowry assay (Lowry *et al.*, 1951) as follows. Samples were made up to 200 ml with distilled water. Lowry solutions A and B were

mixed 50:1 immediately before use and 1 ml of this mixture was added to each protein sample. The sample tubes were vortexed and left at room temperature for 10 minutes. Then 100ml of a 50% (v/v) solution of Folin and Ciocalteu's phenol reagent in water was added to each sample. The sample tubes were vortexed and left at room temperature for 30 minutes before measuring the A_{750} absorbance value. A standard curve of the A_{750} absorbance against protein concentration was plotted using 0, 10, 20, 30, 40 and 50 mg of BSA and the protein concentrations of the samples were determined using this standard curve.

Protein concentration was also determined using the Bradford assay (Bradford, 1976). The dye reagent was purchased from Bio-Rad. 0.25 ml of the reagent was added to 1ml of protein sample diluted in water or buffer. After 10 min the absorbance at 595 nm was recorded against a suitable blank. The reading was compared to a standard curve of BSA standards plotted against absorbance. A fresh standard curve was plotted each time the assay was performed.

2.2.3.4 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Discontinuous Tris-glycine electrophoresis (Laemmli, 1970) was performed on protein samples using Bio-Rad Mini Protean II electrophoresis cells for mini-gels (7 cm x 10 cm) or large gels (20 cm x 20 cm). The polyacrylamide gel consisted of an upper stacking gel of 4.5% acrylamide, 0.12% bis-acrylamide, 0.125 M Tris.HCl, pH6.8, 0.1% SDS (w/v) and a resolving gel of varying concentrations of acrylamide/bis-acrylamide in 0.375 M Tris.HCl, pH8.8, 0.1% SDS (w/v). Gel polymerisation was catalysed by the addition of freshly

prepared ammonium persulphate to a final concentration of 0.1% (w/v) and TEMED to a final concentration of 0.1% (v/v). Gels were placed in a running tank and SDS-PAGE running buffer was added. A fifth of a sample volume of 5 x SDS-PAGE sample buffer was added to each protein sample and the samples were then boiled for 5 min. The prepared protein samples were loaded into the gel wells and run through the gel at 28 mA for about 45 min.

Protein standards were resolved in lanes adjacent to the protein samples to determine their approximate protein molecular masses. SDS7 and Hi-6 molecular weight markers were obtained from Sigma and Rainbow markers were obtained from Amersham.

2.2.3.5 Coomassie staining

Proteins were routinely visualized by staining with Coomassie Brilliant Blue R-250. Gels were successively stained in hot (60°C) Coomassie solutions I, II and III and destained using destain, consisting of 10% glacial acetic acid (v/v), 1% glycerol (v/v). Stained gels were rinsed in distilled water and dried using an Easy Breeze gel drier (Hoefer Scientific Instruments).

2.2.3.6 Silver staining

Silver staining of gels was performed to visualize protein bands of <0.1 μg and is based on the procedure of Heukeshoven and Dernick (1991). All glassware was thoroughly cleaned with household detergent and rinsed in ethanol. Gels were incubated in fixing solution for at least 30min before rinsing in distilled water. They were then placed in

incubation solution for 30min and washed thoroughly 5 times in distilled water for 5 min. Gels were then incubated in silver solution for 40 min and developed in developing solution usually for 5 min. The developing was stopped by replacing the developing solution with stopping solution (2.5 mM EDTA). The gel was preserved in destain solution or in 10% (v/v) glycerol and dried as before.

2.2.3.7 Western blotting

A polyacrylamide gel was soaked in transfer buffer (TTB) for 45 min. Three sheets of Whatman 3MM paper soaked in TBS:Tween 20 were placed onto the anode of an electroblotter, followed by a sheet of nylon membrane and the polyacrylamide gel. Three further sheets of Whatman 3MM paper, which had been soaked in TBS:Tween 20 were placed over the gel. The cathode of the electroblotter was put on top and proteins were transferred from the polyacrylamide gel to the membrane for 3 h at a constant current of 0.8 mA/cm^2 at 4°C. During the transfer the transfer buffer was stirred. Blotting of proteins was briefly confirmed by washing the membranes in Ponceau S stain. The stain was removed by washing the membranes in copious amounts of TBS:Tween 20.

Proteins immobilized on the membranes were then briefly washed in TBS:Tween 20 and soaked in a solution of 10% (w/v) milk powder in TBS:Tween 20 for 30 min with gentle shaking. The membrane was then incubated with primary antibodies diluted in a solution of 5% (w/v) milk powder in TBS:Tween 20 for 30 min. The membrane was washed five times for 5 min in TBS:Tween 20 before incubating with secondary antibodies in a solution of 5% (w/v) milk powder in TBS:Tween 20 for 30 min. The membrane was washed

thoroughly five times in TBS:Tween 20 for 5 min.

Two detection reagents (supplied in a SuperSignal substrate ECL detection kit from Pierce) were mixed in equal proportions, so that a final volume of 0.125 ml per cm^2 of membrane was prepared. This solution was poured over the membrane and left for 5 min. Then the membrane was drained of excess fluid and wrapped in cling film. The protein side was exposed to X-ray film for 15 sec to 10 min and the X-ray film was developed.

2.2.3.8 Fluorescent protease assay using FTC-casein

Proteolytic activity was assayed based on the procedure of Twining (1984). In this procedure a substrate, casein labelled with FITC, is incubated with the test sample. FITC reacts with amino groups of most proteins giving a highly fluorescent fluorescein carbamoyl (FTC) derivative. Proteolytic activity liberates fluorescent TCA-soluble peptides into solution which can be monitored using a fluorimeter.

The assay conditions were as follows unless otherwise stated. For each assay 10 µl of sample was mixed with 20 µl assay buffer and 20 µl 0.5% (w/v) FTC-casein containing 0.2% (w/v) sodium azide. The assay buffer used was 20 mM Tris.HCl, pH7.5, 5 mM $CaCl_2$. This varied from that used by Twining for subtilisin Novo, however in combination with activating calcium ions, the phosphate buffer used in the original assay would have caused a precipitate to form.

Assay samples were incubated at 37°C for up to 18 h. The reaction was stopped by adding 120 µl 5% TCA and mixing. Tubes were allowed to stand at room temperature for 1 h. The TCA-insoluble protein was sedimented at 13,000 x g for 5min. 60 µl of the supernatant was

diluted to 400 µl with 0.5 M Tris.HCl, pH8.5 and mixed. Sample fluorescence was determined using an excitation wavelength of 485 nm and an emission wavelength of 538nm.

2.2.3.9 Protease assay using substrates labelled with methylcoumarin

Two peptide substrates labelled with methylcoumarin (Z-Gly-Gly-Leu-AMC and Boc-Leu-Arg-Arg-AMC.HCl purchased from BACHEM (UK) Ltd.) were used to assess the selectivity of purified Ara12 protease and subtilisin Carlsberg for particular substrate classes.

100 µl assays were set up containing 10 mM peptide substrate (stock prepared in DMSO) and varying concentrations of enzyme sample. The assay buffer used was 20 mM Tris.HCl, pH7.5, 5 mM $CaCl_2$. Proteolytic activity against these tripeptide substrates was monitored at room temperature over time by following an increase in fluorescence. The excitation wavelength used was 355 nm and the emission wavelength was 460 nm.

Chapter 3

Molecular biology of subtilisin-like proteases in *Arabidopsis thaliana*

3.1 Introduction

At the beginning of this project in October 1996, relatively little was known about the molecular biology of subtilisin-like proteases (Slps) in plants in general. In the case of *Arabidopsis thaliana*, one cDNA sequence, *ara12*, had been published and used to investigate tissue expression levels in different tissues (Ribeiro *et al.*, 1995). Nevertheless the role of the Ara12 protease was unknown. Subsequent reports about *slp*s, for example, from tomato (Tornero *et al.*, 1996b) and lily (Taylor *et al.*, 1997), have led to the conclusion that the subtilisin-like proteases are widely found in higher plants and are involved in a number of different processes in plant development and pathogenesis. Although their precise role remains elusive, they may prove to be important modulators of plant growth, possibly acting as part of complex signal transduction pathways.

Moreover, it has emerged that at least some plants, for example tomato plants, appear to simultaneously possess a number of *slp* genes (Tornero *et al.*, 1997). Following this the question was posed as to whether there were any homologues of *ara12* in *Arabidopsis*. This question can be addressed in part by Southern blot hybridisation analysis and by database searches. The former method can be employed to identify very close homologues or slightly more distantly related DNA sequences on the basis of their homologies, depending on the wash conditions used. Database searching has become a vital tool of the molecular biologist, especially so in the last few years. An enormous amount of information is becoming available courtesy of the international effort to sequence the entire genome of *Arabidopsis*. Having retrieved sequence data which could code for Slps, the relevant stretches of DNA can be obtained either by amplifying them by PCR or, providing they exist, by ordering clones from a database stock centre. It is then possible to investigate

the function and mode of action of the proteins coded for by these DNA sequences. Furthermore, if the genomic DNA sequences are known, then this may enable the examination of the gene promoters and give an insight into how they are regulated. Conclusions may be drawn from the presence of known promoter sequence motifs. More concrete evidence can be gained if GUS constructs are generated and transformed into plants. In the constructs the transcription of GUS is regulated by the putative promoter sequences and, with appropriate staining, the precise location of GUS protein can be detected in growing plants. This reveals the tissue expression pattern of the gene of interest. Although it is relevant to mention this here, this subject will be dealt with elsewhere.

The aim of this chapter has been to determine whether there are other genes encoding Slps in *Arabidopsis* in addition to *ara12*. It was proposed to achieve this by Southern blotting and database searching, making particular use of the data which has been generated as part of the *Arabidopsis* genome sequencing project. If there are then it was proposed that these should be amplified by RT-PCR, cloned, sequenced and used to determine the expression levels in *Arabidopsis* tissues by semi-quantitative RT-PCR or Northern blotting. Obtaining cDNA sequences from reverse transcribed plant RNA confirms that genomic DNAs which may be predicted to be genes, using database software, actually give rise to RNA transcripts. Similarly, intron/exon boundaries which have been predicted from genomic DNA sequences could be confirmed using the sequences of the cDNA clones. This information will tell us where in the plant these enzymes are found at the subcellular level and may give a clue as to their functions. As will be discussed, a great deal may also be inferred from comparisons between sequences of homologous plant subtilisin-like

proteases. Striking similarities in amino acid sequences may hint at similarities in protein function. If, for example, a putative *Arabidopsis* Slp displayed a high degree of homology to a tomato Slp known to be involved in microsporogenesis, then the focus of interest in that enzyme would shift and a very specific set of experimental questions can be asked.

Finally, in work done at the University of Durham, the N-terminal sequences of two potential Slps have been reported in tomato extracellular extracts (Robertson *et al.*, 1997). It will be attempted to amplify cDNA sequences coding for these proteases using degenerate oligonucleotide primers designed using the N-terminal sequence data. It is proposed that *de novo* cloning of these *slps* would add to the rather meagre knowledge base concerning this group of plant enzymes, particularly if the genes could be overexpressed. Purification of the gene products would allow the biochemistry of the proteases to be investigated.

3.2 Evidence for *slp* genes in *Arabidopsis thaliana*

3.2.1 Database searches

The amino acid sequence of the Ara12 protease from *Arabidopsis* has been predicted from a cDNA sequence which has been published (Ribeiro *et al.*, 1995). The full-length cDNA sequence of the intronless *ara12* gene has subsequently been deposited in the database (Yamagata *et al.*, 2000) and is shown in Figure 3.1. Before the recent discovery of an auxin-induced cDNA, named *AIR3* (Neuteboom *et al.*, 1999), this protease was the only subtilisin-like protease from *Arabidopsis* about which any information existed in the literature. The Ara12 amino acid sequence has been used to search the National Center for

```
1/1                                                                 31/11                                                               61/21
atg tct tct tcg ttt ctc tcc tcc act gct ttc ttc ctc ctc ctt tgt cta gga ttc tgc cac gtg tcc tct tcc tcc tcc gac caa gga
M   S   S   S   F   L   S   S   T   A   F   F   L   L   L   C   L   G   F   C   H   V   S   S   S   S   S   D   Q   G
91/31                                                               121/41                                                              151/51
act tac atc gtt cac atg gct aaa tct cag atg ccg tcg tct ttc gac ctc cac tct aac tgg tac gat tca tct ctc aga tct atc tcc
T   Y   I   V   H   M   A   K   S   Q   M   P   S   S   F   D   L   H   S   N   W   Y   D   S   S   L   R   S   I   S
181/61                                                              211/71                                                              241/81
gac tca gcc gaa ttg ctc tac act tac gag aac gcg att cat gga ttc tca aca cgg cta act caa gaa gaa gcc gac tcg ctc atg act
D   S   A   E   L   L   Y   T   Y   E   N   A   I   H   G   F   S   T   R   L   T   Q   E   E   A   D   S   L   M   T
271/91                                                              301/101                                                             331/111
caa cct ggt gtt atc tcc gtt tta ccg gag cac cgt tac gag cta cac acc act cgt act cct ctc ttc ctc ggt ctc gac gaa cac acc
Q   P   G   V   I   S   V   L   P   E   H   R   Y   E   L   H   T   T   R   T   P   L   F   L   G   L   D   E   H   T
361/121                                                             391/131                                                             421/141
gca gat ctg ttc cct gaa gcc ggc tct tac agc gac gtc gtc gtc gga gtt ctc gat acc gga gtt tgg ccc gag agt aaa agc tac tcc
A   D   L   F   P   E   A   G   S   Y   S   D   V   V   V   G   V   L   D   T   G   V   W   P   E   S   K   S   Y   S
451/151                                                             481/161                                                             511/171
gac gaa gga ttc ggt cct att cct tcc tcc tgg aaa ggc gga tgc gag gcc gga acc aat ttc acc gct tct ctc tgt aac cgt aaa cta
D   E   G   F   G   P   I   P   S   S   W   K   G   G   C   E   A   G   T   N   F   T   A   S   L   C   N   R   K   L
541/181                                                             571/191                                                             601/201
atc gga gca aga ttc ttc gct cgt ggt tac gaa tca acg atg gga cca atc gat gaa tct aaa gaa tca aga tct cct aga gac gac gac
I   G   A   R   F   F   A   R   G   Y   E   S   T   M   G   P   I   D   E   S   K   E   S   R   S   P   R   D   D   D
631/211                                                             661/221                                                             691/231
ggt cac gga act cac acc tca tcg acc gcc gct gga tcc gtc gtt gaa gga gct agc tta tta ggc tac gct tca gga aca gct cgt ggt
G   H   G   T   H   T   S   S   T   A   A   G   S   V   V   E   G   A   S   L   L   G   Y   A   S   G   T   A   R   G
721/241                                                             751/251                                                             781/261
atg gct cca cgc gct cgt gtc gct gtt tac aaa gtc tgt tgg ctc ggt ggt tgt ttc agc tca gat att tta gct gca atc gat aaa gcc
M   A   P   R   A   R   V   A   V   Y   K   V   C   W   L   G   G   C   F   S   S   D   I   L   A   A   I   D   K   A
811/271                                                             841/281                                                             871/291
atc gcc gat aac gtc aat gtc ctg tcg atg tca ctc ggc ggt gga atg tcg gat tat tat aga gac ggt gtt gcg atc gga gca ttc gcc
I   A   D   N   V   N   V   L   S   M   S   L   G   G   G   M   S   D   Y   Y   R   D   G   V   A   I   G   A   F   A
901/301                                                             931/311                                                             961/321
gcc atg gaa aga ggg att tta gta tct tgc tca gct ggt aat gct ggt ccg agt agt tct agt tta tca aac gta gct cca tgg atc aca
A   M   E   R   G   I   L   V   S   C   S   A   G   N   A   G   P   S   S   S   S   L   S   N   V   A   P   W   I   T
991/331                                                             1021/341                                                            1051/351
act gtt ggt gct ggt act tta gat cgt gat ttt ccg gcg ctt gcg att ctc ggt aac ggc aag aat ttc acc gga gtt tct ttg ttt aaa
T   V   G   A   G   T   L   D   R   D   F   P   A   L   A   I   L   G   N   G   K   N   F   T   G   V   S   L   F   K
1081/361                                                            1111/371                                                            1141/381
gga gaa gct tta ccg gat aaa ttg ttg ccg ttt att tac gct ggg aat gct agt aat gct act aat ggt aat ctc tgt atg acc gga act
G   E   A   L   P   D   K   L   L   P   F   I   Y   A   G   N   A   S   N   A   T   N   G   N   L   C   M   T   G   T
1171/391                                                            1201/401                                                            1231/411
ttg atc ccg gag aaa gta aag ggg aag att gtg atg tgt gat aga gga att aat gct aga gtt cag aaa ggt gat gtg gtt aaa gca gca
L   I   P   E   K   V   K   G   K   I   V   M   C   D   R   G   I   N   A   R   V   Q   K   G   D   V   V   K   A   A
1261/421                                                            1291/431                                                            1321/441
ggt gga gtt gga atg att ctg gct aac act gcg gcg aat ggt gaa gag ctt gtt gcg gat gct cat ttg tta ccg gcg acc acc gtt ggt
G   G   V   G   M   I   L   A   N   T   A   A   N   G   E   E   L   V   A   D   A   H   L   L   P   A   T   T   V   G
1351/451                                                            1381/461                                                            1411/471
gaa aaa gcc ggt gat ata atc cgg cat tac gtg act act gat cct aat ccc acc gct tcg att tca atc tta gga aca gtc gtc ggt gtt
E   K   A   G   D   I   I   R   H   Y   V   T   T   D   P   N   P   T   A   S   I   S   I   L   G   T   V   V   G   V
1441/481                                                            1471/491                                                            1501/501
aaa cca tct ccg gtt gtc gca gcg ttt agc tca cgt gga ccg aat tcg att aca ccg aat att ctt aaa ccg gat ctg atc gct cct gga
K   P   S   P   V   V   A   A   F   S   S   R   G   P   N   S   I   T   P   N   I   L   K   P   D   L   I   A   P   G
1531/511                                                            1561/521                                                            1591/531
gta aac atc ctc gcc gcg tgg acc ggt gct gct gga cca acc gga ctc gct tcc gat tct cgc cgc gtg gag ttc aat atc atc tct ggg
V   N   I   L   A   A   W   T   G   A   A   G   P   T   G   L   A   S   D   S   R   R   V   E   F   N   I   I   S   G
1621/541                                                            1651/551                                                            1681/561
acg tcg atg tct tgc cct cac gtt agt ggt tta gcg gcg ctt ctc aag tcg gtg cat cct gaa tgg agc ccg gcg gcg att aga tcg gcg
T   S   M   S   C   P   H   V   S   G   L   A   A   L   L   K   S   V   H   P   E   W   S   P   A   A   I   R   S   A
1711/571                                                            1741/581                                                            1771/591
ctt atg acc acc gct tac aaa acc tac aaa gac ggt aaa ccg tta ctc gac atc gcg aca ggg aag cct tcg acg ccg ttc gat cac ggt
L   M   T   T   A   Y   K   T   Y   K   D   G   K   P   L   L   D   I   A   T   G   K   P   S   T   P   F   D   H   G
1801/601                                                            1831/611                                                            1861/621
gca gga cac gtg tca cca aca act gcc act aat cca gga ctc atc tac gat cta acg acg gag gat tac tta ggc ttc ctc tgc gca ttg
A   G   H   V   S   P   T   T   A   T   N   P   G   L   I   Y   D   L   T   T   E   D   Y   L   G   F   L   C   A   L
1891/631                                                            1921/641                                                            1951/651
aat tac aca tcg ccg caa att cga agt gtc tcg aga cgt aat tac act tgc gat ccg agt aaa tcg tac tcg gtc gct gat ttg aac tac
N   Y   T   S   P   Q   I   R   S   V   S   R   R   N   Y   T   C   D   P   S   K   S   Y   S   V   A   D   L   N   Y
1981/661                                                            2011/671                                                            2041/681
ccg tcg ttc gcc gtt aac gtt gat gga gtc ggt gcg tat aag tac acg cgc acg gtg acg agc gtg gga gga gct ggg act tac tcg gtt
P   S   F   A   V   N   V   D   G   V   G   A   Y   K   Y   T   R   T   V   T   S   V   G   G   A   G   T   Y   S   V
2071/691                                                            2101/701                                                            2131/711
aaa gta act tcg gag acg aca gga gtc aag att tcg gtt gaa ccg gcg gtt ttg aat ttc aag gaa gct aac gag aaa aaa tcg tat acg
K   V   T   S   E   T   T   G   V   K   I   S   V   E   P   A   V   L   N   F   K   E   A   N   E   K   K   S   Y   T
2161/721                                                            2191/731                                                            2221/741
gtg acg ttt act gta gac tcg tcg aag ccg tct gga tct aac agc ttt ggg agt att gaa tgg tcg gat ggg aaa cac gtg gtg gga agt
V   T   F   T   V   D   S   S   K   P   S   G   S   N   S   F   G   S   I   E   W   S   D   G   K   H   V   V   G   S
2251/751
ccc gtg gcg att agc tgg aca tag
P   V   A   I   S   W   T   *
```

Figure 3.1 Nucleotide sequence of *ara12* (*slpa*) cDNA and it's predicted amino acid sequence. The catalytically important amino acid residues are boxed. A double arrow has been placed above the N-terminal residue of the mature subtilisin-like protease. A near full-length cDNA sequence for *ara12* has been published (Ribeiro *et al.*, 1995). The complete cDNA sequence, as given here, has subsequently been deposited in the database under accession number gb|AF065639 (Yamagata *et al.*, 2000).

Biotechnology Information (NCBI) non-redundant database using the tblastn program, which compares the amino acid sequence against a nucleotide sequence database dynamically translated in all reading frames. The results of this search are shown in Table 3.1 as a list of database accession numbers, along with any designated names, in order of highest to lowest homology with the Ara12 query sequence. Nucleotide sequences which code for known plant subtilisin-like proteases, for example *LIM9* cDNA from lily (Taylor *et al.*, 1997), *ag12* cDNA from alder (Ribeiro *et al.*, 1995) and cucumisin cDNA from melon (Yamagata *et al.*, 1994) have been recognised in this search as well as a large number of sequences which originated from the ongoing sequencing of the *Arabidopsis* genome. cDNA and amino acid sequences have been predicted from the genomic sequences from *Arabidopsis*, and information attached to most of the database entries shown states that these proteins are most likely to be subtilisin-like proteases. According to the results of a search conducted in August 1998 nine putative amino acid sequences were found which showed a distinct similarity to Ara12 and had a particularly high homology around the active site residues. The predicted proteins will be referred to as Slpb-k, reserving Slpa as an alternative name for Ara12. These predicted proteins are marked with asterisks in Table 3.1. *slpe* and *slpf* were initially thought to be two separate genes, but now, according to their database entry, are thought to constitute one gene. All of the predicted proteins are of similar length to Ara12. For database entries where more than one *slp* can be found (this can be seen five times in Table 3.1), the sequences invariably seem to lie immediately adjacent to each other. Where this occurs the entry is marked with a number in brackets indicating the number of adjacent *slp*s present. There is a precedence for this in higher plants, because a genomic cluster has been found in tomato,

Designated name	Accession number	Chromosome number
ara12/slpa *	emb|X85974	5
-	gb|AC005970.2	2
slpk *	dbj|AB010074.1	5
-	dbj|AB022220.1	3
-	emb|AL079347	4
slpb *	emb|AL022023	4
-	dbj|AP000600	3
slpc *	gb|AC002411	1
-	gb|AC020622	1
-	dbj|AB020742	5
XSP1	gb|AF190794	4
-	dbj|AB015475	5
-	dbj|AB012245	5
-(5)	gb|AF118222	4
-(4)	emb|AL049524	4
-	dbj|AB024027	5
-(4)	gb|AC006424	1
-	dbj|AB016890	5
slpe/f *	emb|Z97337	4
-	emb|AL049523	4
-	dbj|AB016885	5
-(2)	emb|AL096859	3
AIR3	gb|AF098632	?
-	gb|AC007178	2
-	emb|AL049171	4
-	emb|AL031187	4
slpd *	gb|AF069299	4
-(3)	emb|AL035527	4
slph *	gb|AF002109	2
-	dbj|AB019229	3
slpg *	gb|AC002392	2
-	gb|AC004708	2
-	dbj|AB017065	5
-	emb|AL078464	4
-	emb|AL080253	4
slpj *	gb|AC003113	1
-	gb|AC007060	1
-	emb|AL132978	3

Table 3.1 Database search results showing potential subtilisin-like protease gene homologues in *Arabidopsis*. The NCBI non-redundant database was searched with the full-length Ara12 amino acid sequence using the tblastn program. For multiple *slp* sequences, found adjacent to each other in the genome, the number of adjacent slp sequences found is shown in parentheses. Three of the sequences referred to have been published: *ara12* (Ribeiro *et al.*, 1995), *AIR3* (Neuteboom *et al.*, 1999) and *XSP1* (Zhao, *et al.*, 2000). Sequences obtained as a result of a search conducted in August 1998 are marked with asterisks. All other sequences listed were first identified in a search in February 2000.

which contains four differentially regulated subtilisin-like protease genes (Jordá *et al.*, 1999). There is evidence that these protease genes may be involved in pathogenesis.

The database search which delivered the *slpb-k* DNA sequences was conducted in August 1998. Subsequently a different *Arabidopsis slp* sequence, *AIR3*, has been cloned from a root cDNA library on the basis that it is transcriptionally upregulated by auxin (Neuteboom *et al.*, 1999). A final database search was performed during February 2000 and led to the results listed in Table 3.1. As before, the non-redundant NCBI database was searched with the full-length Ara12 amino acid sequence using the tblastn program. According to information given on this database, between 70 and 80% of the nucleotide sequence of the *Arabidopsis* genome was accessible at that particular point in time. This revealed evidence for additional *Arabidopsis* subtilisin-like proteases, as might be expected as increasingly more of the plant's genome is sequenced. All the database entries listed in Table 3.1 refer to *Arabidopsis* DNA sequences. Taken together these results would suggest that a large family of subtilisin-like proteases has evolved in *Arabidopsis*. Certainly, this is the case in tomato, as a subtilase gene family has been described in tomato. Currently, 15 members of this family have been cloned (Riggs and Horsch, 1995; Tornero *et al.*, 1996b; Tornero *et al.*, 1997; Meichtry *et al.*, 1999). These subtilase genes have been divided into subfamilies by Meichtry *et al.* (1999) on account of homology between their amino acid sequences. A great deal about gene function can be discovered by looking at sequence homologies. In the same way it may be possible to allude to subfamilies of *Arabidopsis* Slps, if their predicted amino acid sequences are compared against each other. This information may prove invaluable as more is discovered about the function of Slps in plants which seem to have evolved distinct functions (Meichtry *et al.*, 1999).

3.2.2 Southern blot hybridization analysis

Further evidence for the existence of a number of subtilisin-like proteases in this model plant has been found by Southern blot hybridization analysis. After probing a Southern blot of digested genomic DNA it is possible to wash the blot at different stringencies, ie. at different ionic strengths and temperatures. In this way the genomic DNA can be examined using a DNA probe to look for homologues of that sequence. This was done using genomic DNA extracted from the leaves of *Arabidopsis thaliana* ecotype Columbia. The DNA was digested using the different restriction enzymes *Eco*RI, *Hin*dIII, *Bam*HI, *Xba*I and *Sma*I before separating the DNA fragments on a 0.8% agarose gel and transferring them to a Hybond-N nylon filter by Southern blotting. The blot was probed using the full-length *ara12* cDNA, which had been labelled with [α-^{32}P] dCTP using random hexamer primers. The blot was washed at high stringency using 0.1xSSC at 65°C. At this stringency the probe can only bind to genomic DNA sequences showing ~99-100% similarity, i.e. practically identical sequences. The blot was analyzed using a phosphor imaging system (model GS-250 from BioRad) and the hybridization signal image can be seen in Fig.3.2A. The blot was stripped and reprobed with the radiolabelled full-length *ara12* cDNA to look at how a low stringency wash would affect the hybridization. The blot was then washed using 2xSSC at 50°C. At this stringency the probe will be able to bind to genomic DNA sequences showing ~70% similarity or more to the *ara12* probe (Meinkoth and Wahl, 1984). The blot was exposed to X-ray film and the autoradiograph can be seen in Fig.3.2B.

The full-length *ara12* cDNA contains one internal *Eco*RI site, one *Hin*dIII site and one *Bam*HI site, but no *Xba*I or *Sma*I sites. In Fig.3.2A, two hybridization signals can be seen in the *Eco*RI and *Hin*dIII lanes, which would imply that a single copy of the *ara12* gene is

found in the *Arabidopsis* genome. However, as can be seen in Fig.3.2B, a number of bands are visible on the autorad when the blot was washed less stringently. At least ten bands of approximately 2 kb or more can be seen. This suggests that a family of genes closely related to *ara12* exists in *Arabidopsis*.

Taken together, Southern blot analysis and initial database searches conducted in August 1998, strongly suggested that a large subtilase gene family with homology to Ara12 could be found in *Arabidopsis*. A subsequent database search performed in February 2000 revealed a larger number of possible *Arabidopsis* subtilase proteins (approximately fifty). The entire sequence of the *Arabidopsis* genome is now available (Bevan *et al.*, 2001). Therefore, additional putative subtilase genes may be found, which were not identified in either of the previous searches. In order to confirm their presence it was proposed to amplify several of the putative subtilase genes identified by RT-PCR.

3.3 RT-PCR amplification and cloning of *slp*s from *Arabidopsis*

Having identified putative *slp* cDNA sequences via database searches, the next step was to design and synthesize oligonucleotide primers specific for the 5' and 3' ends of the cDNAs to allow their amplification by RT-PCR. Sense and antisense primer pairs were designed and synthesized for this purpose. It was decided to focus on the following seven cDNA sequences: *slpb*, *slpc*, *slpd*, *slpg*, *slph*, *slpj* and *slpk*. The primer pairs used in RT-PCRs are listed in Table 2.2.

It has been possible to amplify several of the corresponding cDNAs via RT-PCR from reverse transcribed total RNA obtained from four day old suspension-cultured *Arabidopsis*

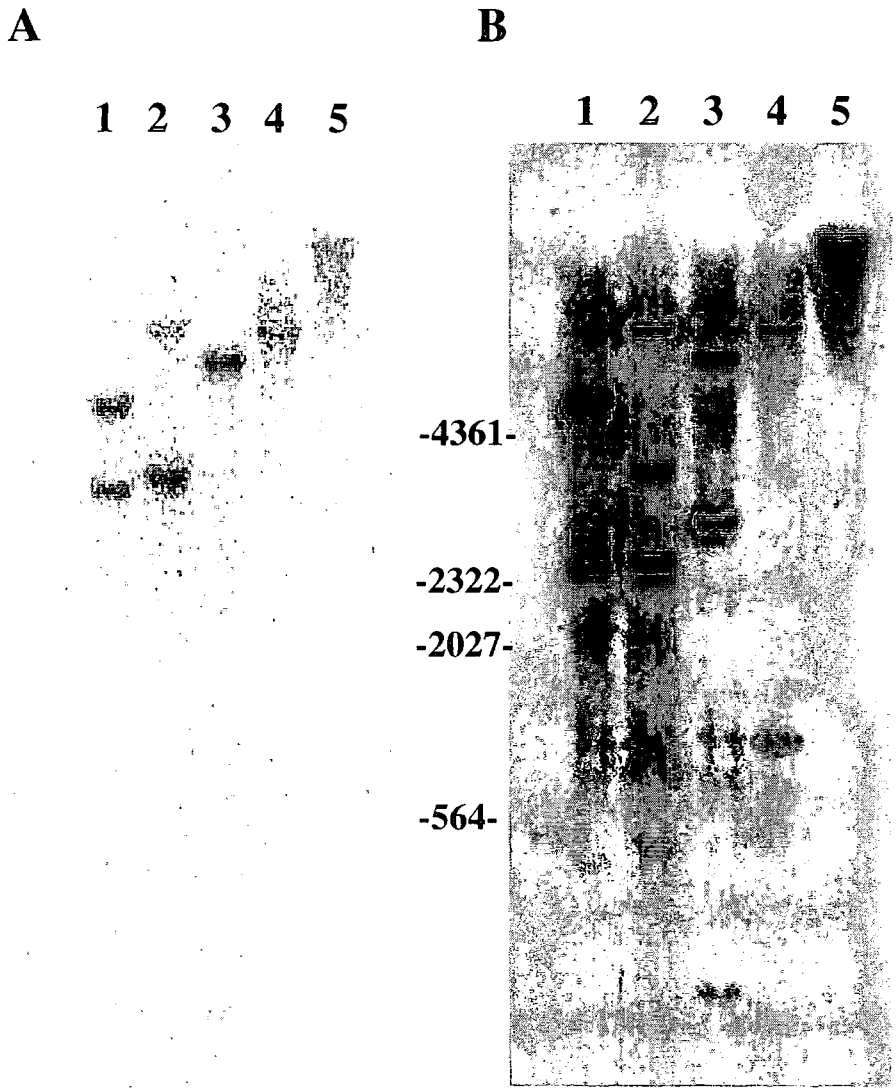


Figure 3.2 Southern blot hybridisation analysis of genomic *Arabidopsis thaliana* DNA using a full-length *ara12* cDNA probe. 6μg of genomic DNA extracted from *Arabidopsis thaliana* Ecotype Columbia leaves was treated with RNaseA and digested with restriction enzymes as indicated below. The DNA fragments were resolved on a 0.8% agarose gel. Lane 1, *Eco*RI; lane 2, *Hin*dIII; lane 3, *Bam*HI; lane 4, *Xba*I; lane 5, *Sma*I. A Southern blot of the genomic DNA was probed with the full-length *ara12* cDNA labelled with ^{32}P dCTP using random hexamer primers. The blot was washed at low stringency (2xSSC, 50°C⇒70% similarity) as shown in **B)** and at high stringency (0.1xSSC, 65°C⇒99% similarity) as shown in **A)**. The size of the DNA fragments is given in base pairs. At least ten bands of approximately 2kb or above can be seen in the *Eco*RI and *Hin*dIII digest lanes of the autoradiograph in **B)**, indicating that there are a number of *ara12* –like subtilase genes in *Arabidopsis*.

cells. Using *Taq* DNA polymerase, PCR products of the expected sizes have been generated in reactions containing the *slpb*, *slpc*, *slpd* and *slpk* primer pairs (see Figure 3.3). However, this was not the case for reactions containing the *slpg*, *slph* and *slpj* primer pairs, although faint bands were visible in the agarose gel shown in Figure 3.4.

The >2 kb amplification products from the *slpb*, *slpc* and *slpk* reactions were gel purified and resolved on agarose gels after digestion using the restriction enzymes *Eco*RI, *Xho*I and *Sma*I (data not shown). The purified *slpb*, *slpc* and *slpk* cDNAs were successfully cloned into the TOPO cloning site of the T vector pCR2.1-TOPO, once confident that their observed restriction pattern was similar to that predicted for these cDNAs.

Two PCR products larger than 2 kb were observed from the *slpd* reaction. One band appeared to be approximately 2.2 kb and the other one was longer than 3 kb. Both DNA bands were gel purified and, although it was attempted to clone both into pCR2.1-TOPO vectors, only the larger of the two was cloned. This was mainly because so little of the 2.2 kb fragment was recovered after gel purification. The best way to clone this smaller fragment would be to reamplify it by PCR, using nested primers if necessary.

The pCR2.1-TOPO constructs generated have been termed pCRslpb, pCRslpc, pCRslpd and pCRslpk on the basis of the insert they contain, as described previously. *Eco*RI restriction digests of these plasmids, containing the four potential *slp* inserts, have been run on 0.8% agarose gels and are shown in Fig.3.5. The approximate sizes of the plasmid inserts have been determined as 1.6 kb for pCRslpb, 2.3 kb for pCRslpc, over 3 kb for pCRslpd and 2 kb for pCRslpk respectively. Nucleotide sequence has been obtained for the5' and 3' regions of all the four cloned *slpb*, *slpc*, *slpd* and *slpk* plasmid inserts using an Applied Biosystems Model 373A DNA sequencer. The sequences which have been deter-

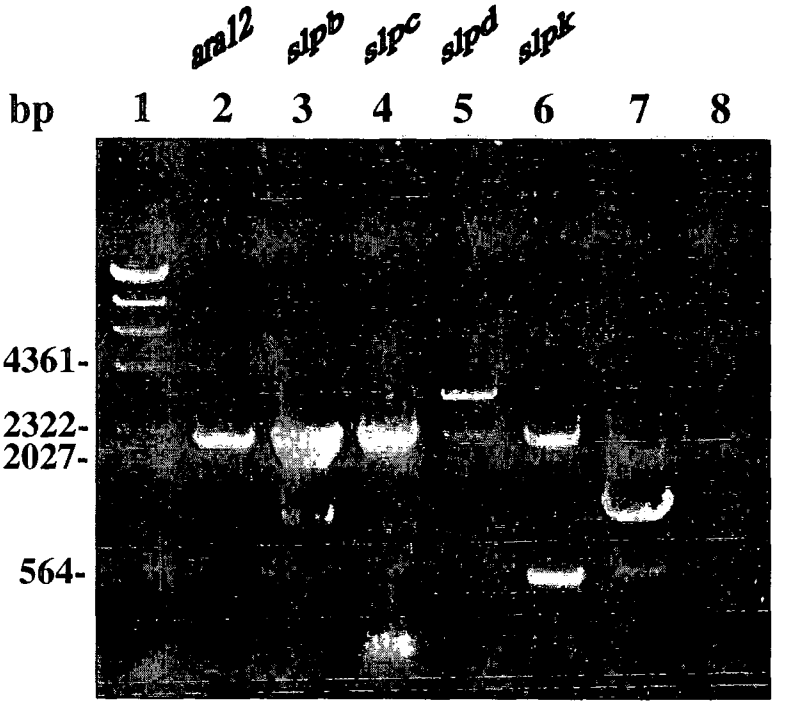


Figure 3.3 RT-PCR amplification of cDNAs (*ara12/slpa*, *slpb*, *slpc*, *slpd* and *slpk*) encoding putative subtilisin-like proteases in *Arabidopsis thaliana* by RT-PCR. Total RNA was prepared from 4 day old *Arabidopsis* suspension cultured cells and reverse transcripts were used in conjuntion with relevant primers to amplify the *slpa*, *slpb*, *slpc*, *slpd* and *slpk* cDNAs. The primers were designed using sequences obtained from database searches described in the text. Each 100µl amplification reaction contained first strand cDNA derived from 2.5µg of total RNA and 10µl from each reaction was run on a 0.8% agarose gel shown here. Lane 1, DNA markers (λ DNA digested with *Hin*dIII); lanes 2-6, RT-PCR products generated using the primer pairs slpa5'/slpa3', slpb5'/slpb3', slpc5'/slpc3', slpd5'/slpd3' and slpk5'/slpk3' repectively; lane 7, RT-PCR with ACT2 primers specific for actin (positive control); lane 8, RT-PCR with no primers (negative control).

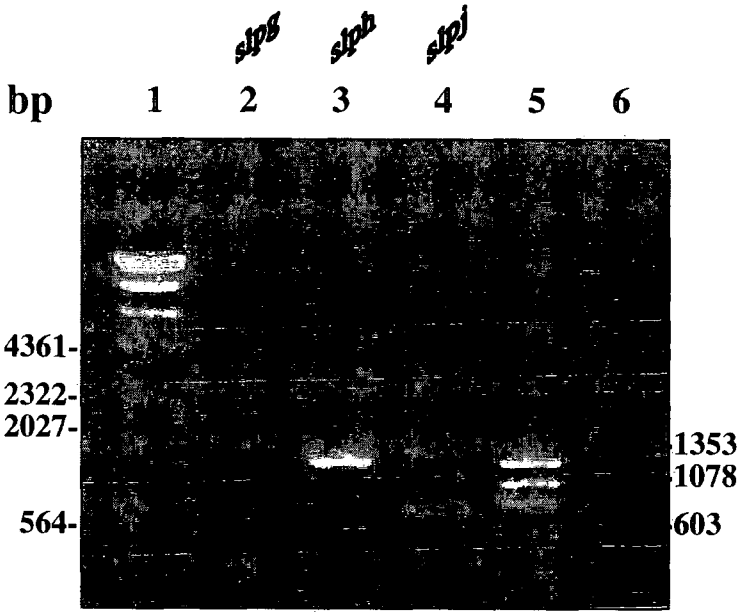


Figure 3.4 Attempted RT-PCR amplification of additional cDNAs (*slpg*, *slph* and *slpj*) encoding putative subtilisin-like proteases in *Arabidopsis thaliana*. Total RNA was prepared from 4 day old *Arabidopsis* suspension cultured cells and reverse transcripts were used in conjuntion with relevant primers to amplify the *slpg*, *slph* and *slpj* cDNAs. The primers were designed using sequences obtained from database searches described in the text. Each 100µl amplification reaction contained first strand cDNA derived from 2.5µg of total RNA and 10µl from each reaction was run on a 0.8% agarose gel shown here. Lane 1, DNA markers (λ DNA digested with *Hin*dIII); lanes 2-4, RT-PCR products generated using the primer pairs slpg5'/slpg3', slph5'/slph3' and slpj5'/slpj3' respectively; lane 5, DNA markers (ϕX174 DNA digested with *Hae*III); lane 6, RT-PCR with no primers (negative control).

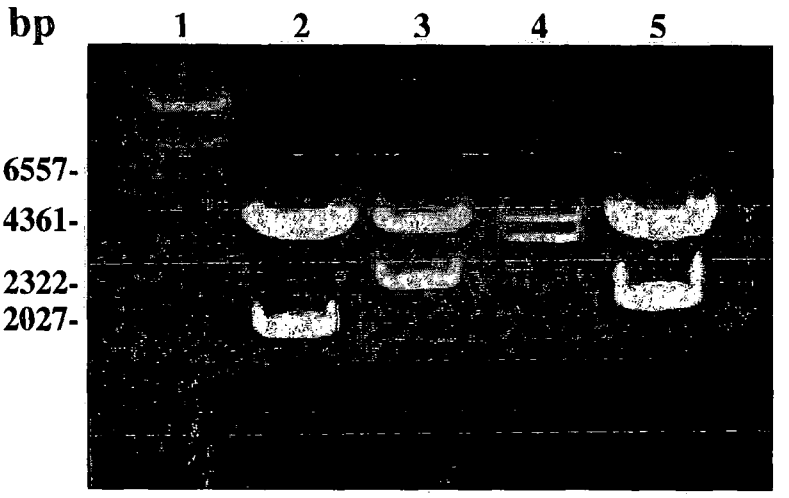


Figure 3.5 *Eco*RI restriction digest of the *slpb*, *slpc*, *slpd* and *slpk* pCR2.1-TOPO clones. Partial *slpb* and *slpk* cDNAs, a full-length *slpc* cDNA and a genomic DNA fragment coding for slpd have been generated by PCR and cloned into the TOPO cloning site of the T vector pCR2.1-TOPO (as confirmed by nucleotide sequencing). Lane1, DNA markers (λ DNA digested with *Hin*dIII); lane 2-5, *Eco*RI digests of the constructs pCRslpb, pCRslpc, pCRslpd and pCRslpk, respectively.

mined in this work are underlined in Figures 3.6-3.8 and 3.10 which show the predicted full-length cDNA sequences as obtained from the database. The lines also demarcate the lengths of the plasmid inserts as they were found in the recombinant plasmids generated.

The sequence data has borne out the fact that the full-length *slpc* cDNA has been cloned, as well as *slpb* and *slpk* cDNA sequences approaching full-length. The *slpd* DNA insert which has been cloned has been shown to be the genomic DNA, corresponding to the ORF of *slpd* cDNA (shown in Figure 3.9). This DNA sequence also contains introns, although many plant *slp* genes have been shown to be intronless. For example, *ara12* is intronless, as are most currently known *slp* genes from tomato (Meichtry *et al.*, 1999).

The question arises at this point as to why the genomic DNA could be cloned even though part of the procedure to generate RT-PCR products involves removal of the genomic DNA from the sample. An outline of the basic RT-PCR procedure used is reiterated below (also see Chapter 2). Total RNA prepared from four day old suspension-cultured *Arabidopsis* cells was treated with RNase-free DNaseI before reverse transcription takes place. This step should remove any residual genomic DNA, which may be present in the RNA preparation. Then first strand cDNA is generated by reverse transcriptase. The solution is heated to inactivate the enzyme and is then ready for use as a PCR template. As a result the template DNA should reflect the mRNA population of the cultured cells and not contain any of their genomic DNA. However, due to the amplification of *slpd* genomic DNA by RT-PCR, it appears that not all of the genomic DNA had been removed in the template used for the reactions shown in Figures 3.3 and 3.4, even though it had been treated with DNaseI. A small amount may have remained, which was sufficient to be amplified over the

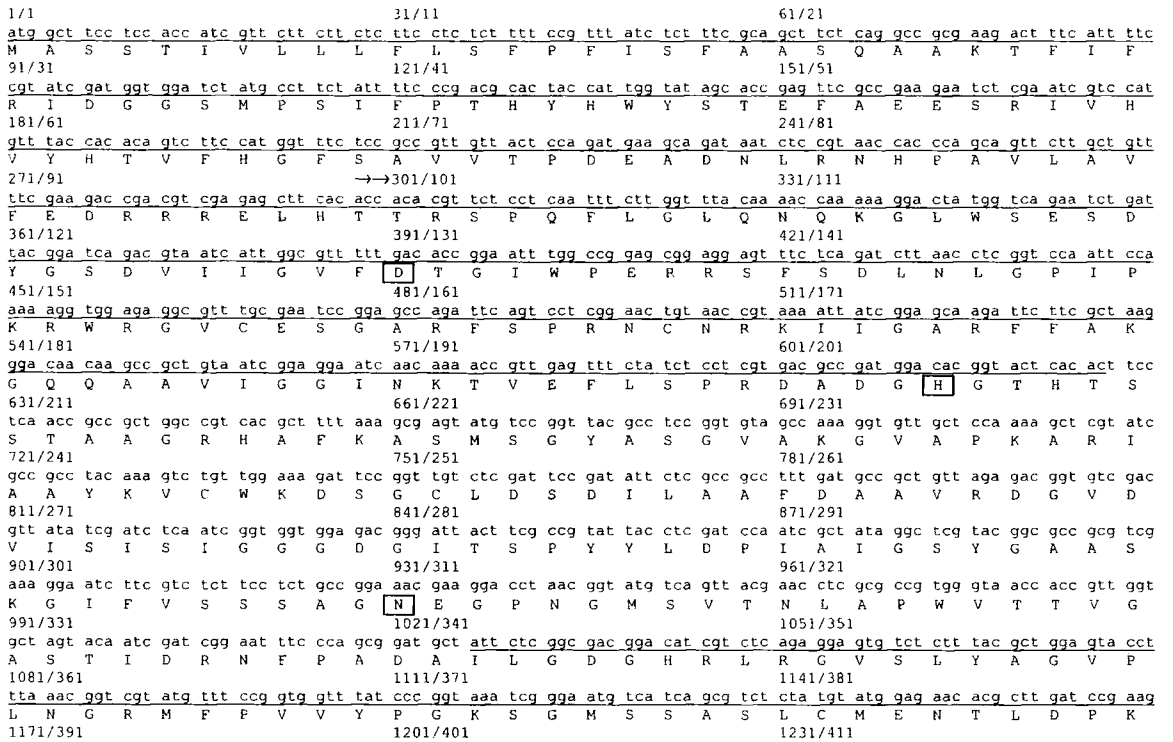


Figure 3.6 Nucleotide sequence of the predicted cDNA *slpb* and its putative amino acid sequence. The catalytically important amino acid residues are boxed. A double arrow has been placed above the N-terminal residue of the mature subtilisin-like protease. Sequences which have been underlined have been sequenced. Underlined sequences demarcate the 5' and 3' ends of the cDNA insert in pCRslpb.

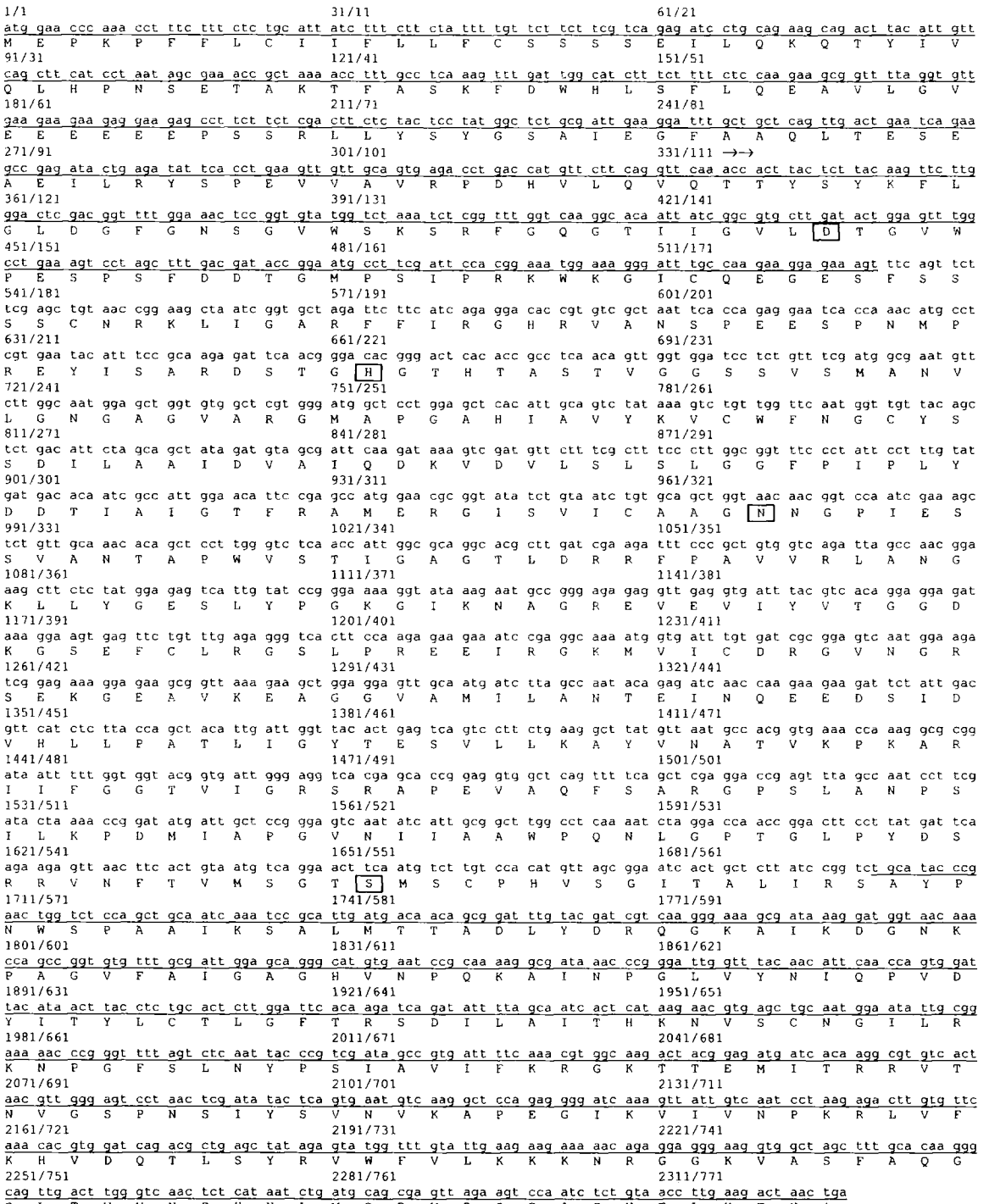


Figure 3.7 Nucleotide sequence of the predicted cDNA *slpc* and its putative amino acid sequence. The catalytically important amino acid residues are boxed. A double arrow has been placed above the N-terminal residue of the mature subtilisin-like protease. Sequences which have been underlined have been sequenced. The full-length slpc cDNA has been cloned in pCRslpc.

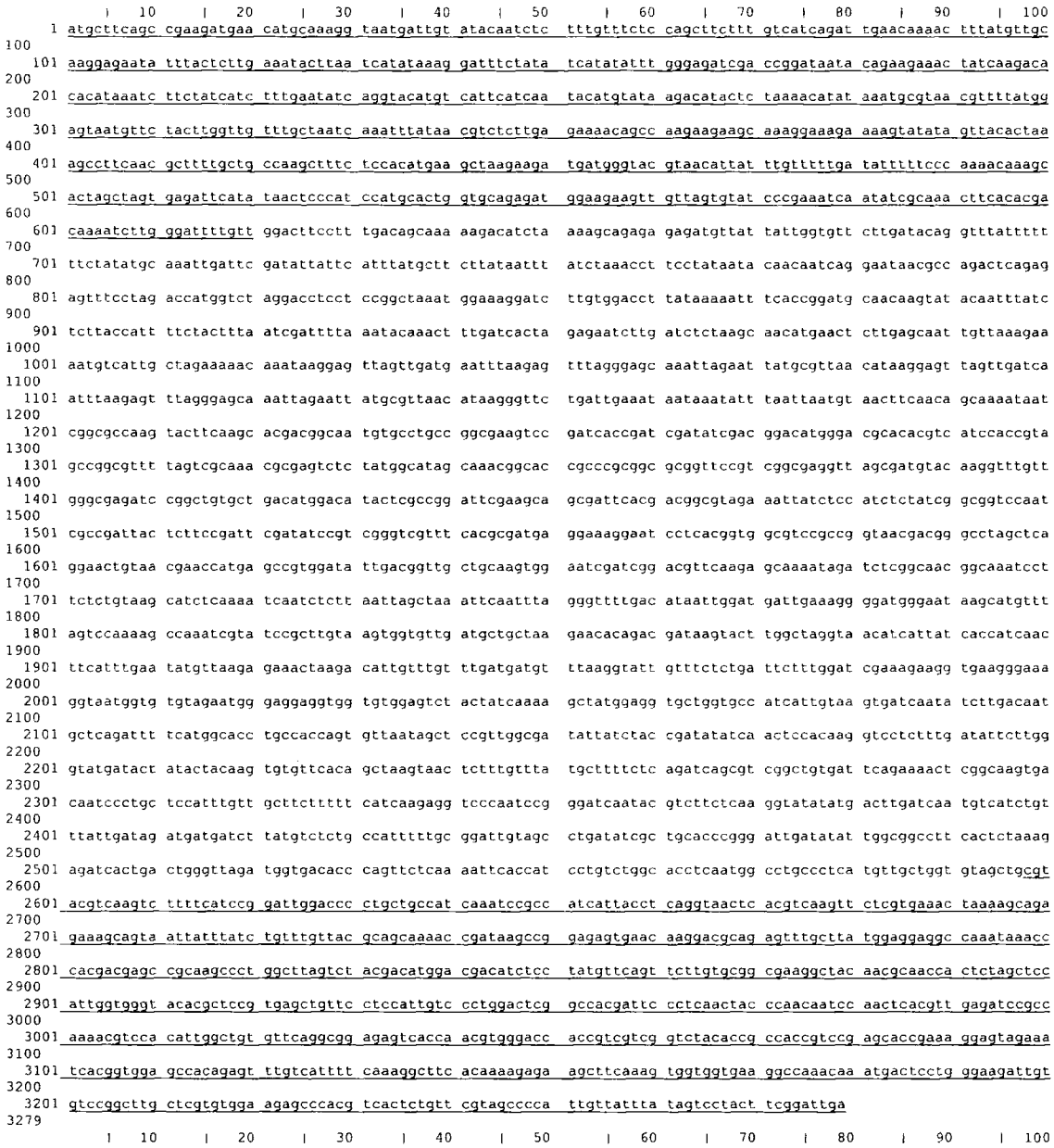


Figure 3.8 Genomic DNA sequence of *slpd* corresponding to the ORF region of the cDNA sequence *slpd*, which is shown in Figure 6.9. Sequences which have been underlined have been sequenced. The genomic sequence shown here has been cloned in pCRslpd.

```
1/1                                                
```

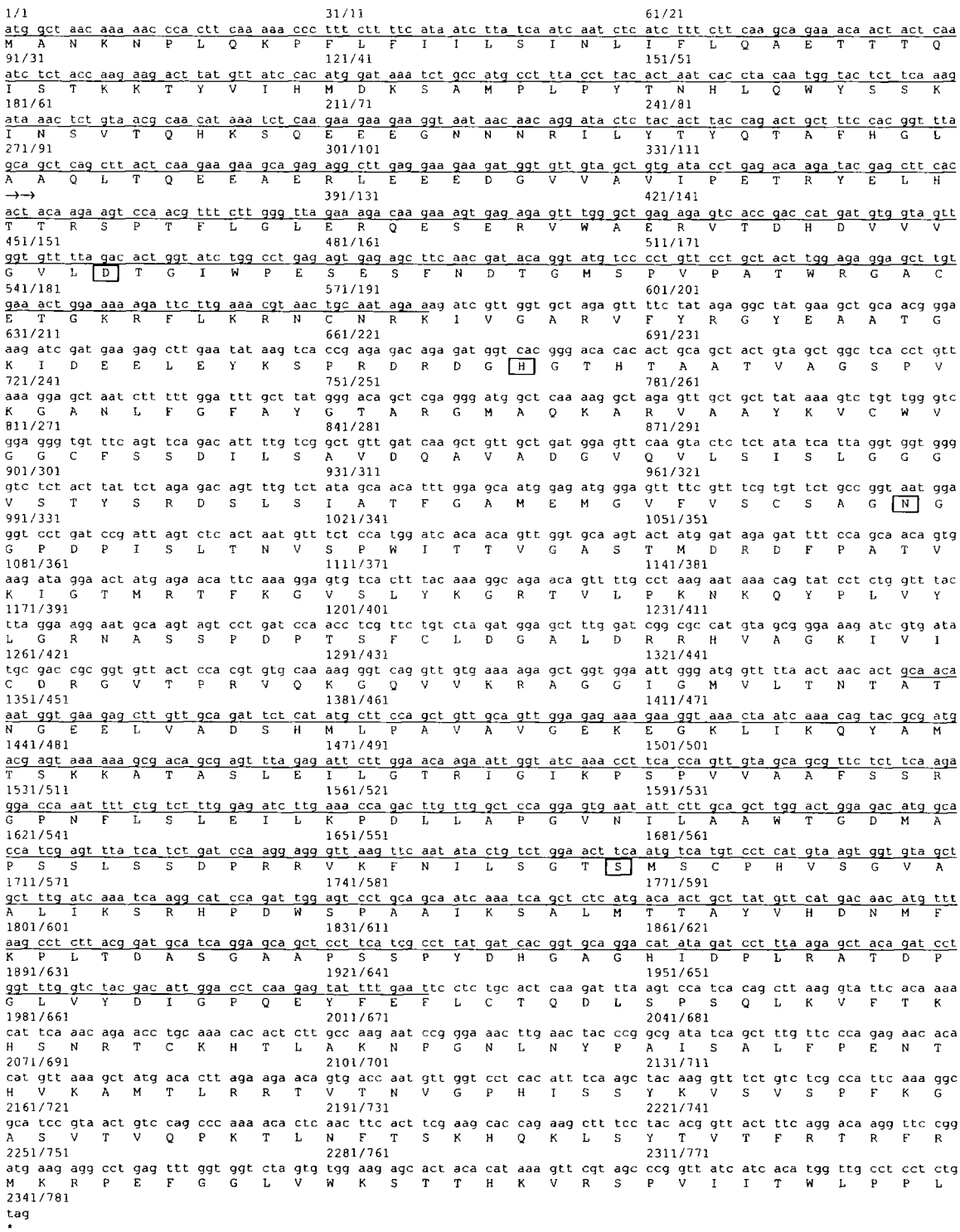


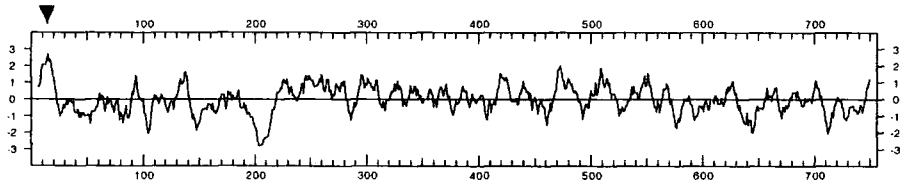
Figure 3.10 Nucleotide sequence of the predicted cDNA *slpk* and its putative amino acid sequence. The catalytically important amino acid residues are boxed. A double arrow has been placed above the putative N-terminal residue of the mature subtilisin-like protease. Sequences which have been underlined have been sequenced. Underlined sequences demarcate the 5' and 3' ends of the cDNA inserts in pCRslpk.

36 cycles of the PCRs and was thus seen on agarose gels. It actually appears as if, in the case of the *slpd* RT-PCR, both the cDNA and the corresponding genomic DNA sequence have been amplified (see Figure 3.3, lane 5). Certainly two DNA species of the correct size to support this idea, have been generated (a 2.2 kb fragment and a >3 kb fragment). It is unclear why the cloned *slpb* and *slpk* cDNA inserts were not quite full-length, even though the primers used in the PCRs should have permitted amplification of full-length sequences. A final extension step of ten minutes should have meant that amplification products were unlikely to remain partial length. However, in a mixed population of varying length cDNAs, the shorter length fragments are preferentially ligated into the vector.

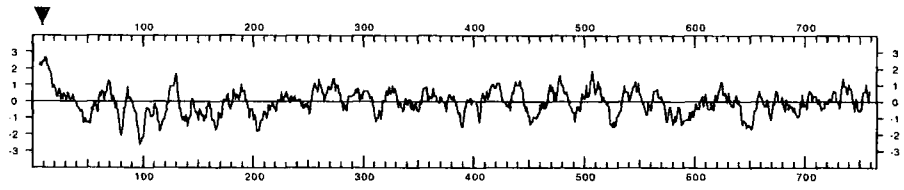
The location of the mature Slp proteins is not known, however information regarding their eventual location can often be obtained using their predicted amino acid sequences. The likely localisation sites of *Arabidopsis* Slps have been predicted using the PSORT program (available on the website http://psort.nibb.ac.jp). Hydrophobic N-terminal signal peptides have been predicted for Slpa, Slpb, Slpc and Slpk, as can be seen in Figure 3.11. These proteases, as well as Slph and the XSP1 protease, have been predicted to be extracellular. The AIR3 protease also has a putative signal peptide necessary for its translocation through the plasma membrane (Neuteboom *et al.*, 1999). As can be seen in Figure 3.11, the Slpd amino acid sequence lacks this hydrophobic peptide region and, like Slpj, is predicted to be found in the cytoplasm or in microbodies. Microbodies are spherical organelles bounded by a single unit membrane and they originate in the endoplasmic reticulum. Microbodies with biologically important roles include peroxisomes and glyoxysomes.

It is interesting to note that the *slp* cDNA sequences which under an identical set of conditions could not be amplified to very high levels (*slpg*, *slph* and *slpj*) are also found to

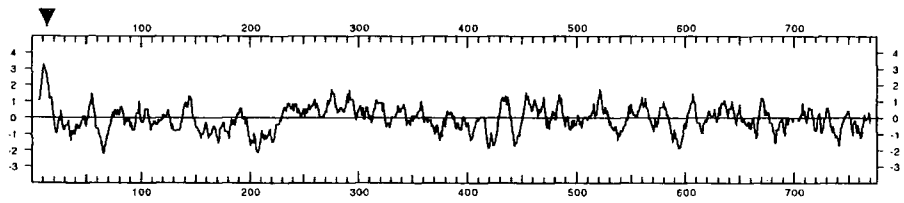
Ara12 (Slpa)



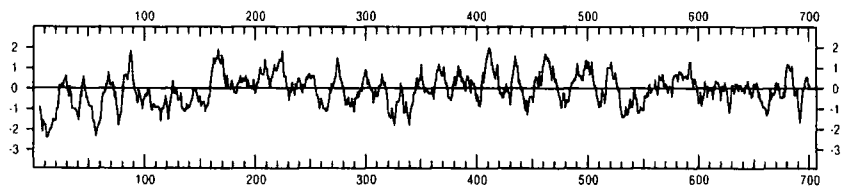
Slpb



Slpc



Slpd



Slpk

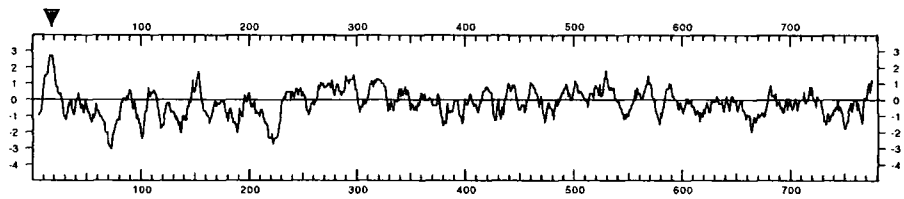
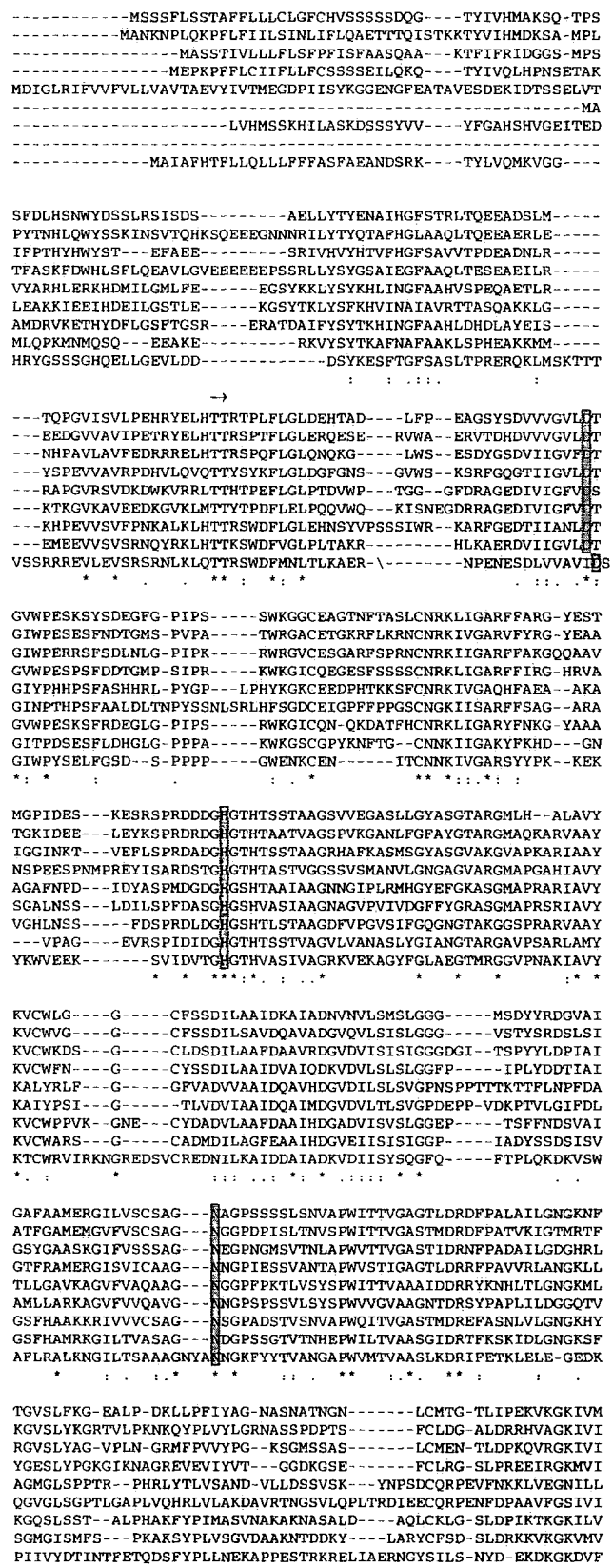


Figure 3.11 Kyte-Doolittle hydropathy plots for the proteins Ara12 (Slpa), Slpb, Slpc, Slpd and Slpk. Four of these five proteins have hydrophobic regions at their N-termini, which are likely to be signal peptides involved in protein targeting. These regions are indicated with arrowheads. Plots were generated using Strider1.3.

be more closely related to each other than to the set of *slp* cDNAs which could be amplified (*slpa*, *slpb*, *slpc*, *slpd* and *slpk*). It is tempting to speculate that the former set of genes may be expressed at lower levels in RNA from four day old suspension-cultured cells than the latter and therefore they cannot be detected by RT-PCR under the conditions of the reaction. A plant subtilase from tomato, *tmp*, appears to be expressed at much lower levels than the other known tomato subtilisin-like protease genes (Meichtry *et al.*, 1999), and is also in a separate subfamily to all other members of the gene family. This protease is known to be involved in microsporogenesis. The tmp protease also shows close homology to the Slpg, Slpj and also, to a lesser extent, to Slph and Slpd (see Figure 3.13). This will be discussed in more detail in Section 3.6. A multiple sequence alignment for the Slp homologues is given in Figure 3.12. Sequence identities and similarities between the Slp homologues Slpa-k have been quantified (see Table 3.2). The phylogenetic relationship between plant subtilisin-like proteases is shown in Figure 3.13.

3.4 Expression of *Arabidopsis* subtilisin-like protease genes

Once the *slp* cDNAs had been isolated it was decided to observe the expression of the corresponding genes in *Arabidopsis* tissues by Northern blot hybridisation analysis. 10μg of total RNA prepared from different *Arabidopsis* tissues (leaf, stem, silique, root and suspension cultured cells) was subjected to electrophoresis through 1% formaldehyde/ agarose gels and blotted onto Hybond-N nylon filters. Each lane in all Northern blots contained an equal amount (10 μg) of total RNA. To confirm that each lane on the blots contained an equal loading, the intensities of the 25S rRNA bands were measured by densitometry using


Slpa/Ara12
Slpk
Slpb
Slpc
Slpg
Slpj
AIR3
Slpd
Slph

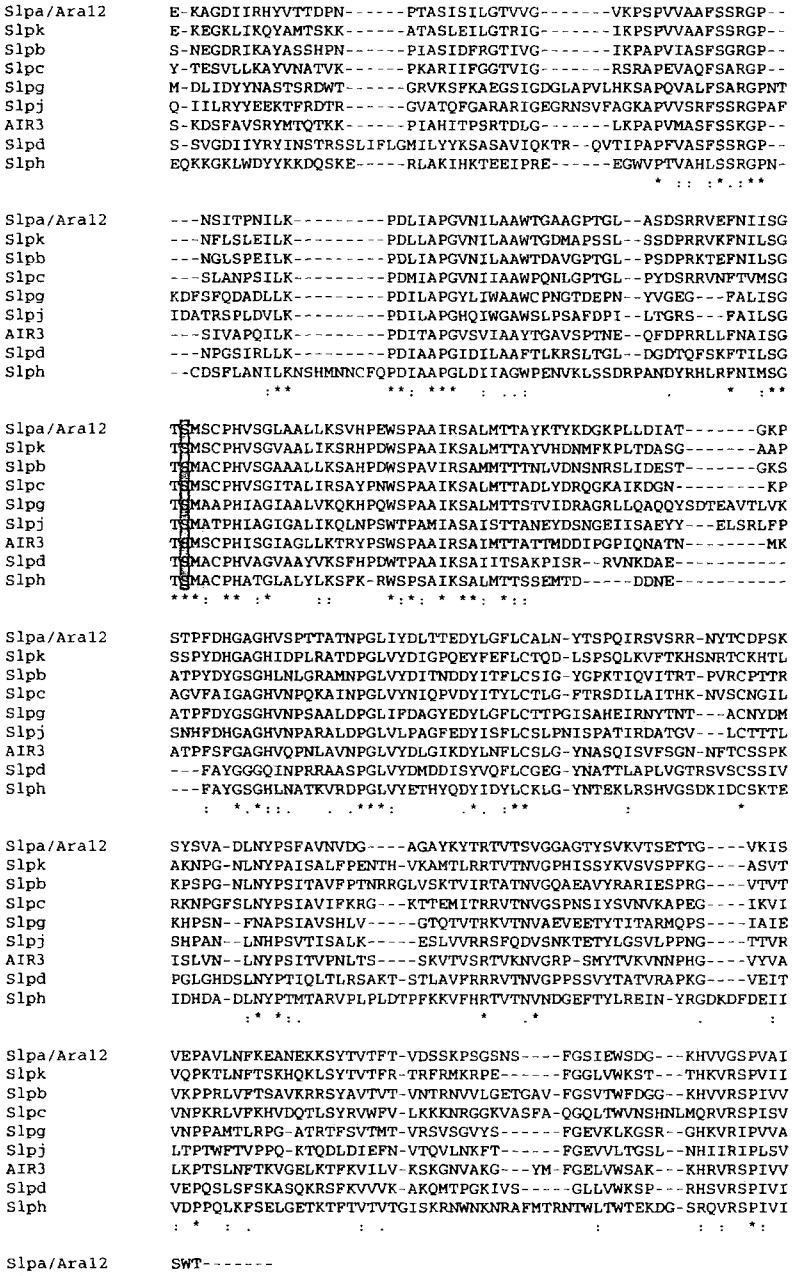


Figure 3.12 Sequence alignment of subtilisin-like proteases from *Arabidopsis*. Deduced amino acid sequences were aligned using ClutalW1.7. Asterisks denote sequence identity, colons denote a strong sequence similarity and dots indicate a weak sequence similarity. The catalytically important amino acid residues are highlighted and the putative N-terminal residues of the mature enzymes have an arrow above them.

	Slpa	Slpb	Slpc	Slpd	Slpg	Slph	Slpj	Slpk
Slpa		47	46	36	32	30	31	51
Slpb	63		45	38	32	29	33	48
Slpc	61	61		36	32	31	31	43
Slpd	52	52	56		31	33	28	37
Slpg	49	46	49	46		27	43	32
Slph	43	41	43	46	41		25	30
Slpj	46	46	46	44	57	39		30
Slpk	65	63	59	52	49	45	48	

sequence identity

sequence similarity

Table 3.2 Comparison of *Arabidopsis* subtilisin-like protease sequences. Amino acid sequences have been predicted from DNA sequences initially found in database searches. cDNA sequences coding for Slpb, Slpc and Slpk, and a genomic sequence coding for Slpd have been cloned. Slpa corresponds to the protein Ara12, which has been described previously. Sequence identities and similarities are given in percentages. The BLAST 2 Sequences program at the NCBI web site was used for pairwise protein-protein sequence comparison (Tatusova and Madden, 1999).

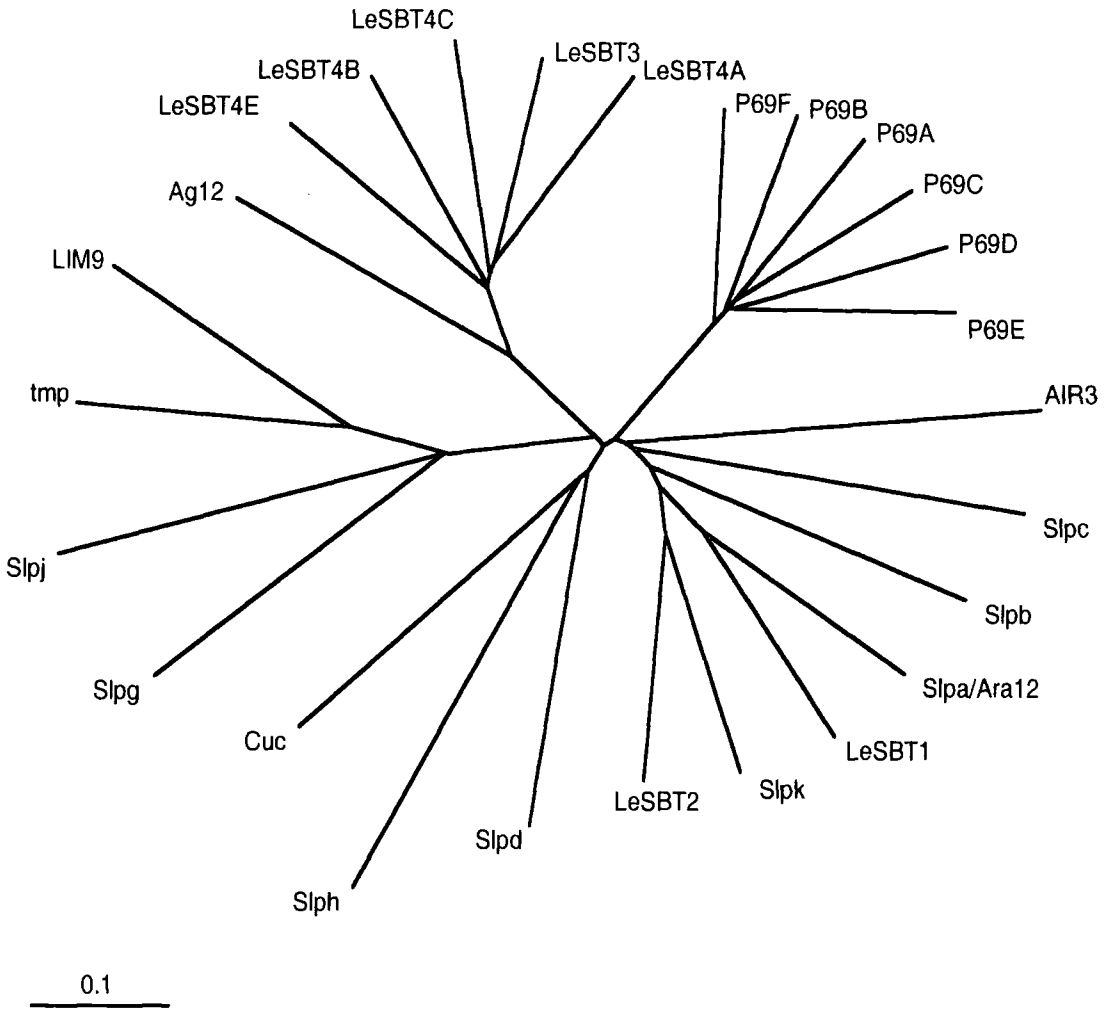


Figure 3.13 Phylogenetic relationship of higher plant subtilisin-like proteases (**Slps**). Amino acid sequences were aligned using the ClustalW1.7 program and the unrooted phylogenetic tree was viewed using TreeView1.5.3. The amino acid sequences of the following proteins shown above have been published: cucumisin from melon (**Cuc**; Yamagata *et al.*,1994), **Ag12** from alder and **Ara12 or Slpa** from *Arabidopsis* (Ribeiro *et al.*, 1995), **LIM9** from lily (Taylor *et al.*, 1997), **AIR3** from *Arabidopsis* (Neuteboom *et al.*, 1999) and **tmp** (Riggs and Horsch, 1995), **P69A** (Tornero *et al.*, 1996b), **P69B** (Tornero *et al.*, 1997), **P69C-F**, **LeSBT1-3** and **LeSBT4A-E** (Meichtry *et al.*, 1999), all from tomato. The **LeSBT4D** sequence from tomato was not included here, because a large part of this cDNA has not been cloned or sequenced. The *Arabidopsis* amino acid sequences **Slpb**, **c**, **d**, **g**, **h**, **j** and **k** were obtained in searches of the NCBI database. Minimal distance between sequences is given in PAM (accepted point mutations per 100 residues).

Quantity One software from BioRad. In addition a separate Northern blot containing the same RNA samples was probed with a cDNA probe homologous to 18S rRNA. This showed that there were equal loadings of RNA on the blot (data not shown).

The full-length *slpc* cDNA and the cloned *slpb* and *slpk* cDNAs were used to probe the Northern blots. For details of the lengths of the cloned slpb and slpk cDNAs see Fig.3.6 and Fig.3.10 respectively. (The 5' two thirds of the predicted full-length *slpb* cDNA and the 5' four fifths of the predicted full-length *slpk* cDNA were used). The cDNAs were radiolabelled using [^{32}P] dCTP and random hexamers. After probing, the blots were repeatedly washed, ending in two stringent washes with 0.1xSSC at 55°C.

The autoradiographs resulting from Northern blots probed with *slpb*, *slpc* and *slpk* cDNAs are shown in Figs.3.14-3.16 respectively. *slpb* mRNA was detected in all tissues being particlarly abundant in silique and leaf tissue, with less being found in root, stem and in suspension-cultured cells (derived from callus from stem tissue). *slpc* transcripts appear to be present in all tissues examined, being particularly abundant in root and stem tissues as well as in the suspension-cultured cells. A smaller amount of *slpc* mRNA was detected in leaf and silique tissues. *slpk* mRNA also appears to be present in all tissues with most found in the stem tissues as well as in suspension-cultured cells. Leaf, root and silique tissues appear to be less abundant in *slpk* transcripts. Slpk appears to be most closely related to Ara12 as can be seen in the phylogenetic tree shown in Figure 3.13. Slpb and Slpc are also slightly more distantly related to Ara12, however more closely so than the other putative subtilisin-like proteases Slpd and Slpg-j.

It has not been attempted to quantify the relative levels of expression by densitometry,

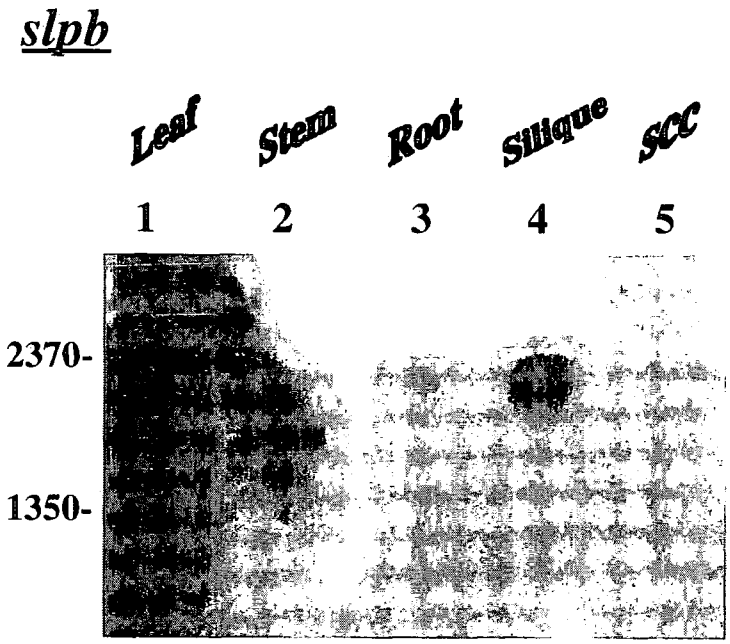


Figure 3.14 Expression of *slpb* mRNA as determined by Northern blot hybridisation analysis. Total RNA was prepared from *Arabidopsis* leaf, stem, root and silique tissues and from suspension-cultured cells (SCC). A Northern blot (10μg for each sample) was probed using the *slpb* cDNA probe.

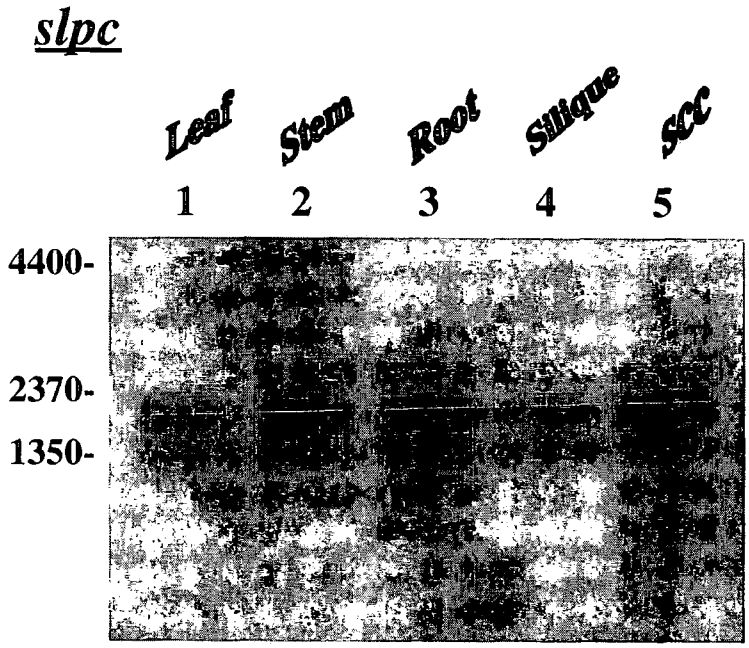


Figure 3.15 Expression of *slpc* mRNA as determined by Northern blot hybridisation analysis. Total RNA was prepared from *Arabidopsis* leaf, stem, root and silique tissues and from suspension-cultured cells (SCC). A Northern blot (10μg for each sample) was probed using a full-length *slpc* cDNA probe.

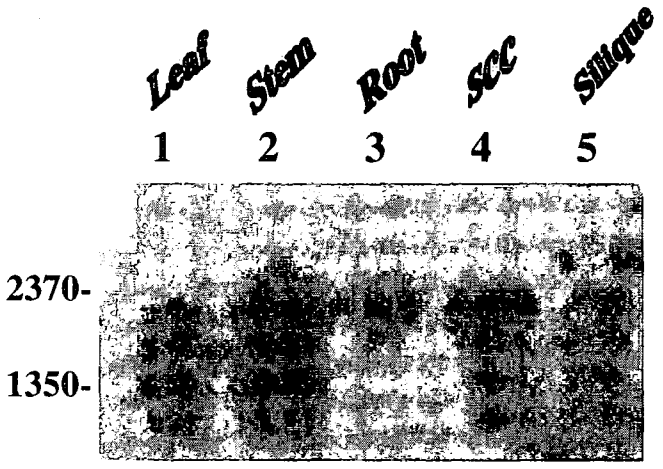


Figure 3.16 Expression of *slpk* mRNA as determined by Northern blot hybridisation analysis. Total RNA was prepared from *Arabidopsis* leaf, stem, root and silique tissues and from suspension-cultured cells (SCC). A Northern blot (10μg for each sample) was probed using the *slpk* cDNA probe.

DNA probe	Leaf	Stem	Root	Silique	Suspension-cultured cells
slpb	XX	X	X	XX	X
slpc	X	XX	XX	X	XX
slpk	X	XX	X	X	XX

Table 3.3 Summary of the relative abundances of *slp* transcripts in *Arabidopsis* tissues as determined by Northern blot hybridisation analysis. Northern blots containing 10μg of total RNA prepared from the tissues stated were probed using *slpb*, *slpc* or *slpk* cDNAs as described. **Key:** **X**=transcript found at relatively low levels; **XX**=transcript found in relative abundance.

because the experiments would really need to be repeated several times, to build up a more accurate picture, before this should be done. Instead a table has been drawn up which summarizes the findings in a more qualitative way, enabling the levels of each *slp* transcript to be compared for different tissues (see Table 3.3). Ideally, the Northern blot analysis should be repeated several times to be confident of the relative expression levels of Ara12 in different tissues.

3.5 *De novo* cloning of DNA encoding a tomato subtilisin-like protease

From previous work at the University of Durham (Robertson *et al.*, 1997), two N-terminal sequences bearing a resemblance to other plant subtilisin-like protease sequences were determined from primary cell wall extracts of tomato suspension culture cells (see Table 3.4). Both these proteins were found in sodium chloride extracts and had masses of 76 kDa and 80 kDa. In the paper the 18 amino acid residue N-terminal sequence (accession number P80815) of the 76 kDa protein is reported as sharing 50% sequence similarity to another tomato subtilisin-like protease. In the course of this work it was noticed that the previous entry (accession number P80814) also has a striking resemblance to features associated with subtilisin-like proteases.

The 76 kDa and 80 kDa proteins were, for the purpose of the work described here, called STOM1 (for subtilisin from tomato) and STOM2 respectively. It was attempted to amplify DNA coding for these proteins by PCR. To achieve this aim it was necessary to use degenerate oligonucleotide primers to prime the PCRs. Genomic DNA used in this exercise

was prepared from cell suspension cultures derived from the same tomato cultures used in the initial study (Robertson *et al.*, 1997).

Two degenerate primers, STOM1F and STOM2F, were designed using the two N-terminal sequences. A third degenerate oligonucleotide was designed using a region of the protein highly conserved amongst all known plant subtilisin-like proteases. Probably the most useful region, in terms of the high degree of sequence conservation is the sequence GHGTHTA, which was used in the design of this third primer. Inosine was included at positions of high degeneracy (of 3 or 4), because it can hybridize to any nucleotide. This greatly reduced the degeneracies of the resulting primers. The sequences of the degenerate primers synthesized are listed in Table 2.2, as well as being shown in Table 3.4 with the relevant amino acid sequences used in their design. A number of factors had to be taken into account when designing the degenerate DNA primers, including the length of the primer, the degeneracy and the annealing temperature range (range of temperatures over which the mixture of oligonucleotides will anneal to the corresponding template DNA). It is important to consider this last variable, because otherwise the primer may not be able to bind to the template DNA sequence, resulting in failure of the amplification reaction.

Having synthesized the primers and calculated the annealing temperature range, the amplification reactions were carried out using a robocycler with a differential annealing temperature heating block. This meant that samples could be cycled at 94°C for 1 minute, followed by 44-66°C (with 2°C increments between samples) for 1 minute, followed by 1.5 minutes at 72°C. Samples containing tomato genomic DNA and either the primer combination STOM1F/ STOM2R or STOM2F/ STOM2R were initially denatured at 94°C

Extraction buffer	Band	Accession number	N-terminal sequence	M_r	Sequence similarity
1M NaCl	A	P80814 (STOM2)	STHTSDFLKL	80kDa	none given
1M NaCl	C	P80815 (STOM1)	STRTPEFLGLDNQCGVWA	76kDa	tomato subtilisin protease (50%)

Designed oligonucleotide primers and their degeneracies

STOM1F (16-fold)

```
5'-GGAATTC-AC/TAAC/TCAA/GTGC/TGGIGTITGGGC-3'
  - EcoRI -D   N    Q    C    G  V  W  A -
```

STOM2F (256-fold)

```
5'-GGAATTC-CAC/TACIA/TC/GIGAC/TTTC/TC/TTIAAA/GC/TT-3'
  - EcoRI -H    T  S      D    F    L    K    L   -
```

STOM1R (16-fold)

```
5'-GGAATTC-GCIGTA/GTGIGTICCA/GTGA/C/G/TCC-3'
  - EcoRI - A  T    H  T  G    H        G-
```

Table 3.4 Design of degenerate oligonucleotide primers specific for two identified tomato subtilisin-like proteases. N-terminal amino acid sequences, obtained from tomato cell suspension culture extracts, which resemble subtilisin-like proteases are shown with their accession numbers (data taken from Robertson *et al.*, 1997). Both the proteins sequenced were found in the NaCl non-sequential primary cell wall extract. The degenerate primers STOM1F and STOM2F designed using these N-terminal sequences are shown, along with STOM1R another degenerate primer designed using a highly conserved plant subtilisin-like protease sequence (GHGTHTA). Primer degeneracies are shown in brackets.

for 5 minutes, incubated as described above for 40 cycles, and then there was a final extension step of 10 minutes at 72°C.

A small aliquot of each of the samples was then resolved on 2% agarose gels. Apart from primer dimers, no DNA bands were visible with any of the STOM1F/STOM2R samples. A DNA band of between 300 and 350 bp was visible with those STOM2F/STOM2R samples which had been cycled with an annealing step between 44-52°C. A 2% agarose gel showing just these five samples can be seen in Figure 3.17. No PCR amplification was observed when the annealing temperature was raised. The size of the band generated correlates perfectly with the expected size for a fragment of DNA, coding for the stretch of plant subtilisin-like protease under examination. After gel purification, this DNA fragment was cloned into the TOPO cloning site of the T vector pCR2.1-TOPO. An *Eco*RI digest of the recombinant plasmid generated is shown in Figure 3.17. The insert has been sequenced in both directions and the amino acid sequence has been predicted. This DNA sequence and a comparison of the predicted amino acid sequence to another extracellular subtilisin-like protease, Ara12, is shown in Figure 3.18.

After obtaining a *STOM2* DNA fragment, it was proposed to use it to isolate the corresponding full-length cDNA sequence by screening a library. However, a paper has subsequently been published (Meichtry *et al.*, 1999) containing predicted amino acid sequences for twelve subtilases from tomato. The STOM2 sequence was found to be identical to part of one of these sequences, namely *LeSBT4D*. A full-length *LeSBT4D* cDNA is not available to the authors, as their sequence is truncated due to the insertion of a retrotransposon-like element. In view of these published results this work was not taken

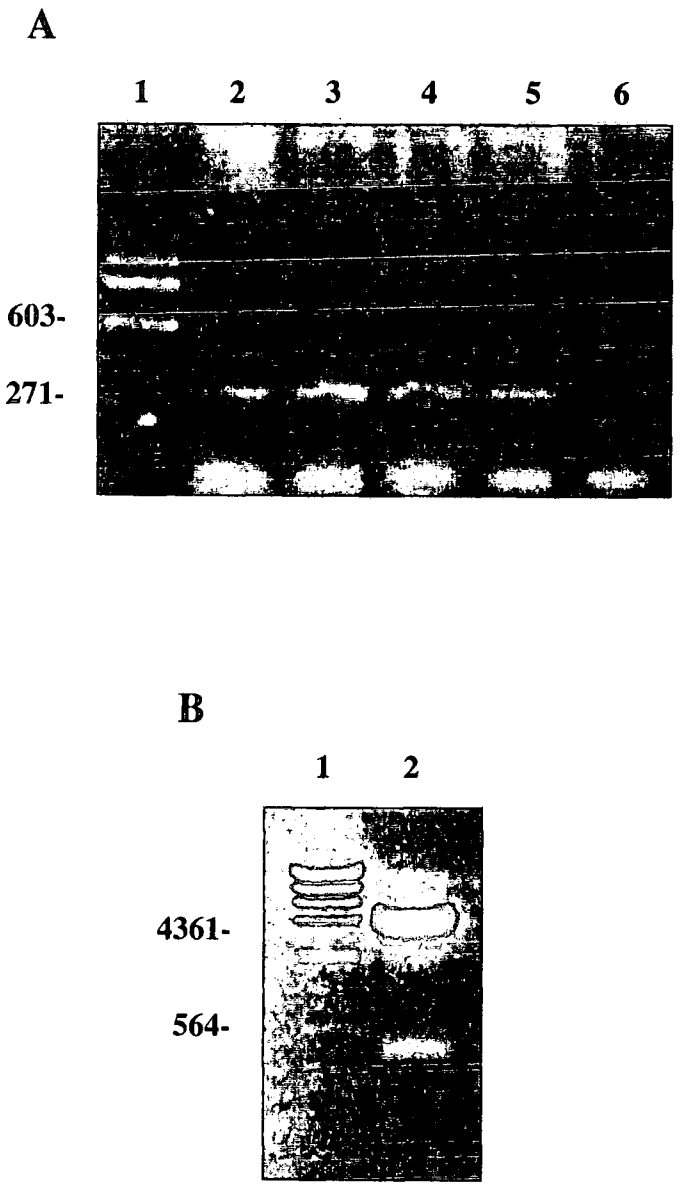


Figure 3.17 Cloning of part of a tomato subtilisin-like protease gene. **A)** PCR amplification of a *STOM2* DNA fragment using the STOM2F/STOM2R primer combination at different annealing temperatures. Lane 1, DNA markers (ϕX174 DNA digested with *Hae*III), lanes 2-6, PCR with the STOM2F/STOM2R primer combination using the annealing temperatures 44°C, 46°C, 48°C, 50° and 52°C, respectively. **B)** Cloning of the amplified *STOM2* PCR product (shown in A) into the pCR2.1-TOPO vector. Lane 1, DNA markers (λ DNA digested with *Hin*dIII), lane 2, *Eco*RI digest of the recombinant pCR2.1-TOPO vector containing the *STOM2* insert. DNA marker sizes are given in base pairs.

A

```
1/1                                     31/11
CAC ACG AGG GAT TTT TTG AAG CTT ACT CCT TCG TCG GGG TTG TGG CAA GCT TCT GGT TTA
H   T   R   D   F   L   K   L   T   P   S   S   G   L   W   Q   A   S   G   L
61/21                                   91/31
GGA CAA GAT GTG ATC ATT GGG GTT CTT GAC GGT GGA ATC TGG CCA GAA TCC GCG AGT TTC
G   Q   D   V   I   I   G   V   L   D   G   G   I   W   P   E   S   A   S   F
121/41                                  151/51
AGA GAT GAT GGT ATG CCT GAA ATA CCC AAA AGG TGG AAA GGT ATT TGC AAG CCA GGC ACT
R   D   D   G   M   P   E   I   P   K   R   W   K   G   I   C   K   P   G   T
181/61                                  211/71
CAG TTT AAT ACT TCA ATG TGC AAC AGA AAA CTG ATT GGG GCT AAT TAC TTC AAT AAG GGA
Q   F   N   T   S   M   C   N   R   K   L   I   G   A   N   Y   F   N   K   G
241/81                                  271/91
ATT TTG GCG AAT AAT CCT ACT GTG AAA ATC TCC ATG AAT TCT GCC AGG GAT ATT GAT GGT
I   L   A   N   N   P   T   V   K   I   S   M   N   S   A   R   D   I   D   G
301/101
CAT GGC ACC CAC ACC GC
H   G   T   H   T
```

B

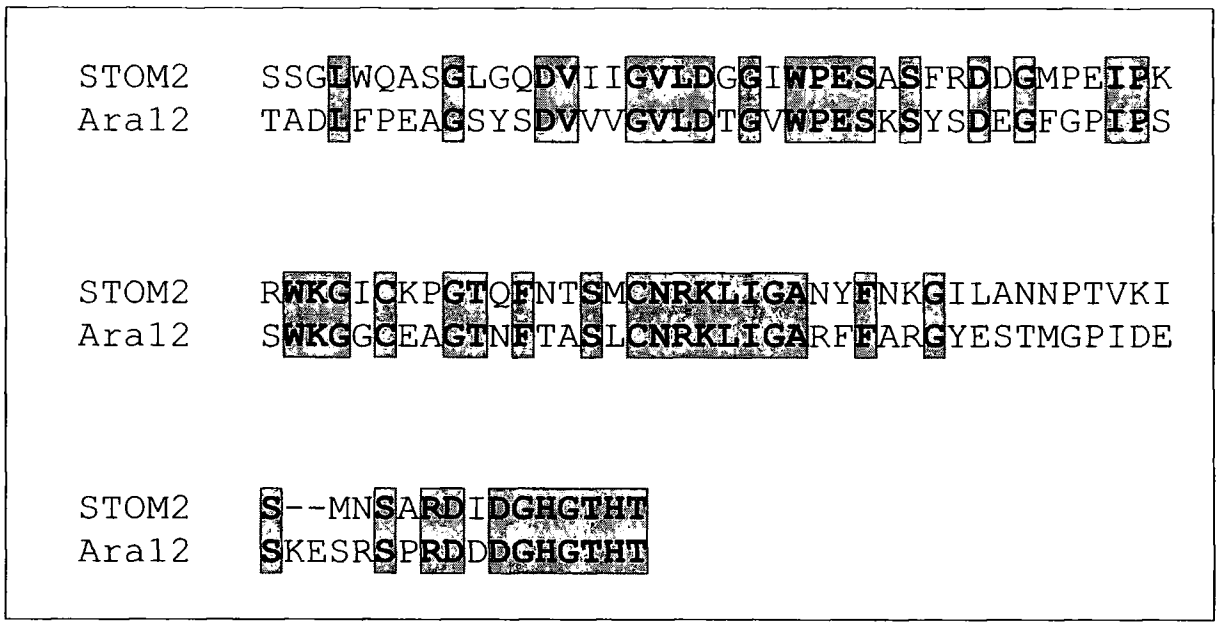


Figure 3.18 A) Nucleotide sequence of the *STOM2* DNA fragment and the predicted amino acid sequence. **B)** Comparison of the predicted amino acid sequence of STOM2 with the amino acid sequence of Ara12. The predicted STOM2 sequence shown has 48.5% identity (47/97) to the corresponding region of Ara12 at the amino acid level. Identical residues are highlighted. The sequences were aligned using the ClustalW program.

any further. Surprisingly, the gene or cDNA coding for the 76 kDa STOM1 protein still remains uncloned.

What has emerged from this and previous work (Robertson *et al.*, 1997) is that an extracellular location can be confirmed for two identified tomato subtilisin-like proteases. Part of the cDNA coding for one of these proteins, STOM2/LeSBT4D, has been cloned here and elsewhere (Meichtry *et al.*, 1999). In the latter paper a putative processing site is indicated on the amino acid sequences, where a speculative propeptide region is cleaved to reveal the mature enzyme. Protein sequencing (Robertson et al., 1997) verifies the position of the N-terminal amino acid residue of the mature enzyme, as well as its location in the extracellular matrix.

3.6 Conclusions

The existence of a subtilisin-like protease gene family in *Arabidopsis thaliana* has been demonstrated by Southern blot hybridization analysis. Database searching has borne this out and shown that there may be more than fifty or more Slps in *Arabidopsis* alone. Several DNA sequences coding for putative Slps (*slpb*, *slpc*, *slpd* and *slpk*) have been amplified by RT-PCR and cloned. Sequence data obtained from these clones has confirmed intron/exon boundaries which have been predicted and placed in the NCBI database. The attempted cloning of a number of other *slp* cDNAs (*slpg*, *slph* and *slpj*) has also been described.

A parallel to this has been seen with the discovery of a host of protein kinases in *Arabidopsis* during the last decade (Hardie, 1999). The precise function of most of these protein kinases is still unknown, however work is progressing to find their protein

substrates and their probable roles in signal transduction. Inferences in this task can be drawn from what is known about protein kinases and protein phosphorylation in other eukaryotic organisms, including budding yeast, for example (Hanks *et al.*, 1988).

Northern blot hybridization analysis has been used to show the tissue expression of three related *Arabidopsis* subtilase genes, *slpb*, *slpc* and *slpk*. These genes appear to be expressed in all tissues to different degrees. Whilst *slpb* mRNA was found predominantly in the leaf and silique, *slpc* mRNA was found in stem and root tissue and *slpk* transcripts were found mainly in the stem tissue. According to the PSORT program, which predicts protein localization sites, the mature Slpa, Slpb, Slpc and Slpk enzymes will be found in the extracellular matrix. The mature Slpa protease is known to be secreted into the medium of suspension-cultured cells and, thus in plants it is indeed assumed to be localized in the extracellular matrix.

Also a *de novo* cloning approach to clone two previously uncharacterized tomato Slps has been described. The cloning of part of an extracellular subtilisin-like protease gene has been successfully achieved.

A multiple sequence alignment has been generated using deduced *Arabidopsis* amino acid sequences and their sequence identities and similarities have been quantified by pairwise comparison. The phylogenetic relationships between known plant subtilisin-like proteases have been examined and are displayed in Figure 3.13. From the clustering observed in the tree, a case can be made for the presence of subfamilies within the family of plant subtilisin-like proteases. Several potential subfamilies can be identified, but the important question to bear in mind is whether this can offer an insight into the properties and function of novel proteases. Although minimal reference to the actual function of the *Arabidopsis*

Slp enzymes has been made in this chapter, some ideas concerning this will be put forward in Chapter 7, formed on the basis of sequence homologies between these related plant proteases.

Chapter 4

Generation of polyclonal antibodies against and immunolocalisation of Ara12/Slpa subtilisin-like protease

4.1 Introduction

A number of plant subtilisin-like proteases have been purified and an increasing amount of information concerning the DNA sequences coding for these proteases is becoming available. However the tissue specificity of expression is known for relatively few of these proteases and, more specifically, there are only two plant subtilisin-like proteases for which a precise cellular or subcellular localisation has been established. One is LIM9 protease from lily, and using immunocytochemistry it has been found to be secreted into the anther locule, particularly in the tapetal cells (Taylor *et al.*, 1997). It is thought to be involved in microsporogenesis. This is the only plant subtilisin-like protease, and there may be more than fifty of these enzymes in *Arabidopsis* alone, for which an immunolocalisation has been reported. The precise cellular location of a second plant subtilisin-like protease, Ag12, has been elucidated by *in situ* hybridisations of root nodule sections with ^{35}S-labelled RNAs (Ribeiro *et al.*, 1995). The protease was found exclusively in root nodules of alder trees which have been infected with *Frankia* bacteria.

The aim of this chapter has been to generate polyclonal antibodies which specifically recognize the Ara12/Slpa protease in *Arabidopsis*. The expression of the *ara12* gene has already been established by Northern blot hybridisation analysis (Ribeiro *et al.*, 1995). However, once generated it was envisaged to use these antibodies to confirm, by Western blotting, the tissue specificity of this enzyme at the protein level. The presence of gene transcripts does not necessarily mean that transcript translation has occurred, nor does it indicate the presence of an active protease. It was further proposed to investigate where this protease was located at the cellular and subcellular level, using secondary antibodies labelled with gold particles. These particles can be visualized in the cell by electron

microscopy. This should reveal a great deal about the role of this enzyme in the plant, especially in conjunction with studies proposed in this work into the nature of the substrate specificity. If for example this protease is capable of degrading most proteins it comes into contact with, it would be particularly important to find out exactly where in the plant the enzyme is located. This information might disclose a great deal about the function of this subtilisin-like protease.

Initially it was intended to overexpress *ara12* cDNA encoding the protease in *E. coli* using an appropriate plasmid construct. This overexpressed protein would be used to raise antibodies and to characterise the protease biochemically, for example to determine its substrate specificity. As the work progressed it was realized that there was a problem in generating a construct encoding the full-length mature protease, and so a partial length construct, attached to the *malE* gene from *E. coli* was generated. The *malE* gene was used in the expression system, because its product, maltose binding protein (MBP), was necessary for an affinity purification of the resulting fusion protein. The fusion protein was overexpressed and used to raise antibodies in rabbits. Antisera were used to investigate the tissue specificity of the protease as well as its cellular and subcellular location. The antisera were also utilized, to verify the identity of the enzyme, in the initial stages of the purification of the Ara12/Slpa protease, which was undertaken using *Arabidopsis* cell suspension cultures. The purification of Ara12/Slpa protease has been described in Chapter 5.

4.2 Overexpression and purification of Ara12/Slpa antigen

In order to obtain Ara12 antigen to generate polyclonal antibodies specific for Ara12 protease, it was proposed to overexpress all or part of the *ara12* cDNA in *E. coli* cultures using expression constructs. The overexpressed protein could then be recovered, purified and used to raise antibodies. This would be simple if the protein of interest had been tagged, as affinity chromatography could be employed. Initially it was attempted to clone the *ara12* cDNA encoding the full-length mature protease between the *Nde*I and *Not*I cloning sites of the expression vector pET24a. A histidine tag on the C-terminus of the overexpressed protein in this system is used to subsequently purify it from crude cell extracts. First strand cDNA was generated from total RNA prepared from *Arabidopsis* suspension cultured cells, using reverse transcriptase. Vent polymerase was used to generate *ara12* cDNA by PCR using primers SUB278 and SUB2239 listed in Table 2.2. The PCR product was digested with *Nde*I and *Not*I and set up with suitably prepared pET24a (which had been digested with the same restriction enzymes and phosphatase treated), in ligation reactions.

There appeared to be a genuine problem in cloning this cDNA into the vector. No viable transformants could be generated. A colony hybridisation screen of one of the attempts at cloning the cDNA in-frame into the cloning cassette of pET24a, revealed the presence of several positive bacterial colonies. A number of black dots were observed on the corresponding autoradiograph, suggesting the presence of colonies containing the desired recombinant plasmid (data not shown). But when the appropriate colonies were located, they had an unusual morphology. The colonies appeared to be thinner and more transparent than normal and cover a slightly larger area than expected. They looked as if they been

dissolved. It was not possible to culture these colonies in LB medium or on fresh LB agar plates.

If the promoter of the expression construct was "leaky" and transcription occurred without prior induction (addition of IPTG), then an active protease may have been synthesized in the cells during early stages of growth of the bacterial colonies. This would have caused lysis of the transformant colonies.

To review this problem it was proposed to insert *ara12* cDNAs into a blunt ended cloning site of an expression vector. From these results it could be deduced whether there was a bias for the insert to be introduced in the sense or the antisense orientation. A bias in generating antisense orientated constructs, but no sense orientated constructs could indicate that low transcription levels of *ara12* were occurring and causing death of bacterial cells containing the sense construct.

In order to ascertain this experimentally, it was necessary to use a second expression vector, because pET24a does not contain any blunt-ended cloning sites from which protein expression can occur. The expression vector pMAL-c2 was chosen, as it contains the blunt-ended cloning site *Xmn*I. *ara12* cDNA was prepared by RT-PCR from total RNA from *Arabidopsis*, with the primer pair MAL239/MAL2247 (listed in Table 2.2), using Vent polymerase, so that it was blunt-ended. Inserting the cDNA insert prepared into the pMAL-c2 *Xmn*I site in the sense orientation would result in the cDNA being in-frame and correct transcription of *ara12* could occur. In theory 50% of the recombinant plasmids should have the insert in the sense orientation.

PCR analysis of the resulting eighteen transformant bacterial colonies (using the *ara12* primer MAL239 and the pMAL-c2 specific vector, pMALFOR, listed in Table 2.2) proved

that all of the recombinant plasmids generated contained the *ara12* cDNA in the antisense orientation. No sense constructs had been generated. This was assumed to be because transcription (and ultimate translation) of *ara12* had resulted in cell death. Other factors may have accounted for these results, as anecdotal evidence suggests that a 50:50 ratio of sense:antisense insertion rarely occurs and is highly dependent on the DNA species being ligated.

To obtain antigen to raise antibodies, it was proposed to use the pMAL system (New England Biolabs.) to overexpress a C-terminal portion of Ara12 as a fusion protein bound to maltose binding protein (MBP). The portion of Ara12 used in this fusion protein lacked the active site residues, up to and including the active site serine residue, and so would have no protease activity. No problems of cell lysis would then be possible, but a C-terminal protein (216 amino acid residues in length) would be synthesized for use in antiserum production. An outline of the pMAL protein fusion and purification system can be seen in Figure 4.3, showing the overexpression and purification of part of the Ara12 protease.

A 650bp DNA fragment coding for the relevant C-terminal section of the Ara12 enzyme was amplified from first strand cDNA prepared from total RNA from *Arabidopsis* suspension cultured cells. Vent polymerase was used to give blunt-ended PCR products. The 650bp amplification product, shown in Figure 4.1A, was gel purified and ligated into the *Xmn*I site of pMAL-c2. In-frame cloning of the insert in the sense orientation was confirmed by PCR analysis, using the primer combination pMALFOR and MAL2247, and nucleotide sequencing. The resulting recombinant plasmid created was named pMAL-ara12 and is shown in Figure 4.1B.

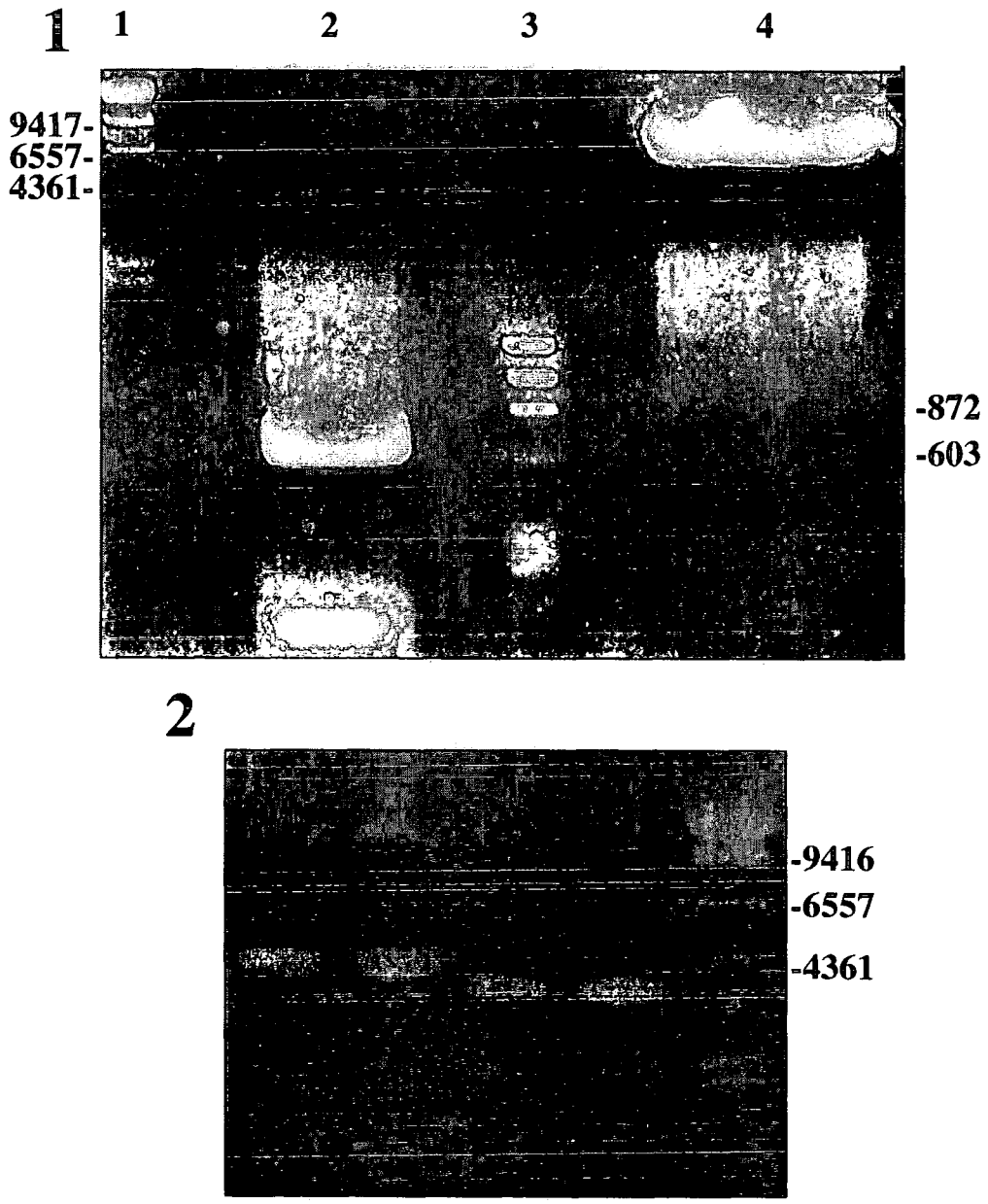


Figure 4.1A Construction of the overexpression construct pMAL-ara12. **Panel 1** shows the *ara12* fragment amplified by RT-PCR using Vent polymerase and cloned into pMAL-c2. Lane 1, λ *Hin*dIII DNA markers; lane 2, the 650bp DNA fragment, coding for a C-terminal part of Ara12; lane 3, an *Xmn*I digest of pMAL-c2; lane 4, ϕX174 *Hae* III DNA markers. The DNA fragment in lane 2 was ligated into the *Xmn*I site of pMAL-c2. **Panel 2** shows unrestricted pMAL plasmids; the first two lanes on the left hand side show the pMAL-ara12 construct and the next two lanes show pMAL-c2 for comparison. From this it can be seen that pMAL-ara12 is larger than pMAL-c2, consistent with the insertion of a 650bp DNA fragment.The lane on the right hand side shows λ *Hin*dIII DNA markers. A map of the overexpression construct is shown in Figure 4.1B.

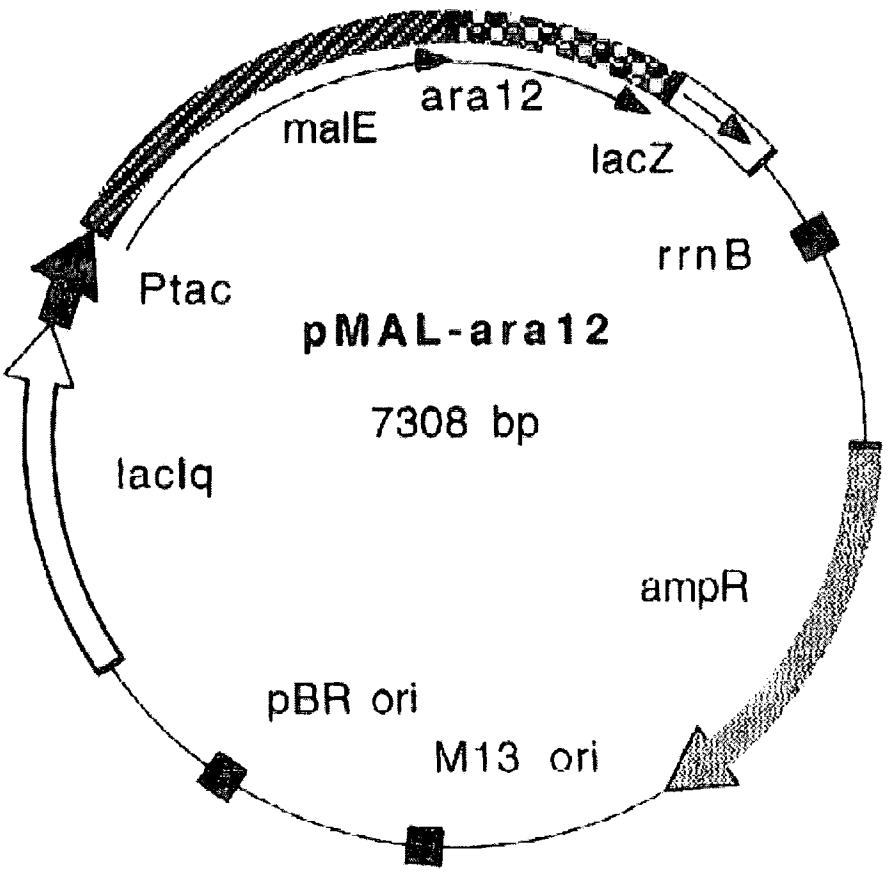


Figure 4.1B Map of the overexpression construct pMAL-ara12 which was generated by inserting *ara12* cDNA (encoding the C-terminal 216 amino acid residues) into the *Xmn*I cloning site of the pMAL-c2 plasmid. It contains the inducible P_{tac} promoter to transcribe the *malE-ara12* gene fusion. The *lacI*^{q} gene codes for the Lac repressor, which turns off transcription from P_{tac} until IPTG is added. Ampicillin resistance (ampR) is afforded by the β-lactamase gene. A portion of the rrnB operon containing two terminators prevents transcription originating from P_{tac} from interfering with plasmid functions.

To assess overexpression of the fusion protein, the pMAL-ara12 plasmid construct was transformed into DH5α, XL1 Blue and BL21 DE3 *E. coli* cells which had been made competent using calcium chloride. An induction experiment was performed to determine which cells produced (most) fusion protein. Samples were taken every hour for three hours after IPTG induction of the cells and, after boiling to promote solubilization, were run on 10% gels to determine the protein profile (see Figure 4.2). The same was done for equivalent cells which had not been induced with IPTG. When compared to the untreated cell cultures, the IPTG treated cultures showed that a protein was generated which was slightly smaller than the 66 kDa molecular weight protein marker. The expected size of the fusion protein was approximately 65 kDa, made up of the MBP portion (42 kDa) and the Ara12 portion (estimated to be 23 kDa). The implications of these results were that the MBP-Ara12 fusion protein could be overexpressed in *E. coli* cells and that XL1 Blue cells overexpressed the fusion protein to particularly high levels. Therefore fusion protein was synthesized on a larger scale in XL1 Blue cells as described below, before proceeding with its affinity purification. See Figure 4.3 for an overview of this procedure.

One litre of LB media, supplemented with glucose to a final concentration of 10 mM and ampicillin at 50 μg/ml was inoculated with an overnight culture derived from twenty fresh XL1 Blue transformant colonies. The cells were incubated at 37°C, shaking at 200 rpm, until they reached an OD_{600} of 0.7 (approximately 12 hours). IPTG was then added to the cultures to a final concentration of 0.4 mM. Three hours after induction, the *E. coli* cells were pelleted by centrifugation at 5,000 x g at 4°C for 15 minutes. The pellet was resuspended in 50 mls of ice cold column buffer (10 mM Tris.HCl, pH7.4, 200 mM NaCl, 1 mM EDTA, 1 mM sodium acetate) and passed twice through a French press to lyse the

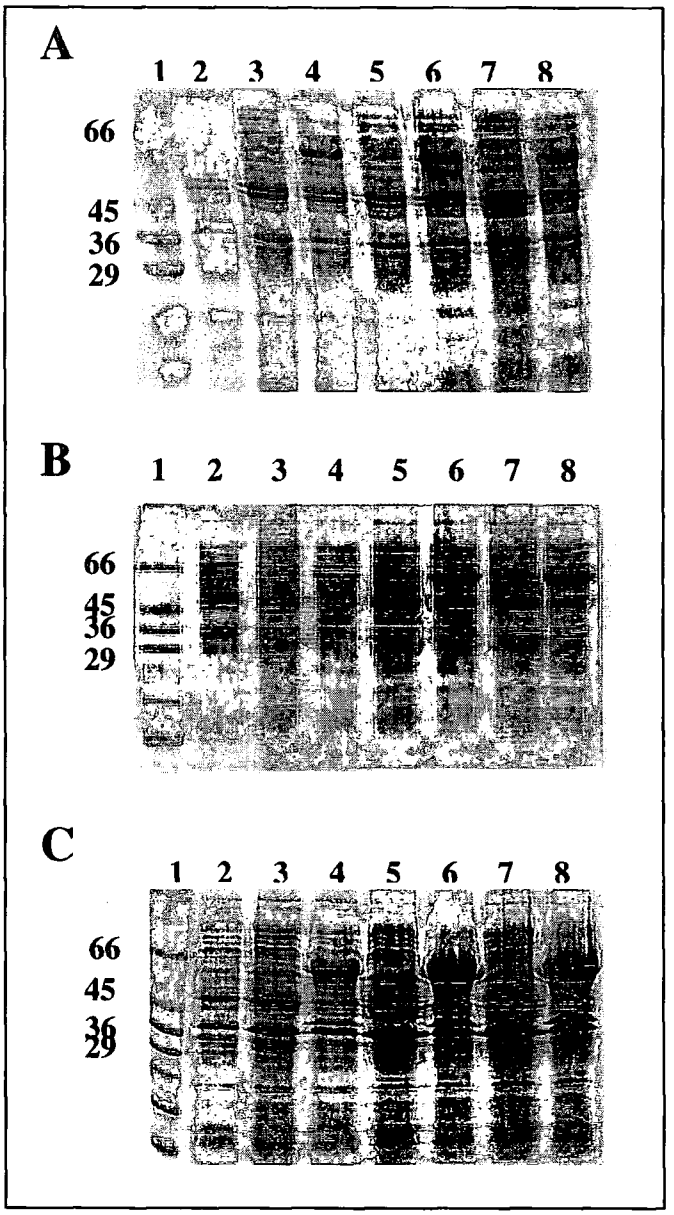


Figure 4.2 Protein profile of *E. coli* cell cultures transformed with the pMAL-ara12 plasmid after induction with IPTG. This shows the overexpression of the MBP-Ara12 C-terminal fusion protein. Each of the three panels shown represents the protein profile from different bacterial strains used: **A)** DH5α, **B)** BL21 DE3, **C)** XL1 Blue. For each bacterial strain the same lane loadings apply, as follows: lane 1, SDS7 protein molecular weight markers; lanes 2, 3, 5 and 7, non-induced cells harvested 0, 1, 2 and 3 hours after other cells were induced; lanes 4, 6 and 8, cells harvested 1, 2 and 3 hours after induction with IPTG. Molecular weights have been given in kilodaltons.

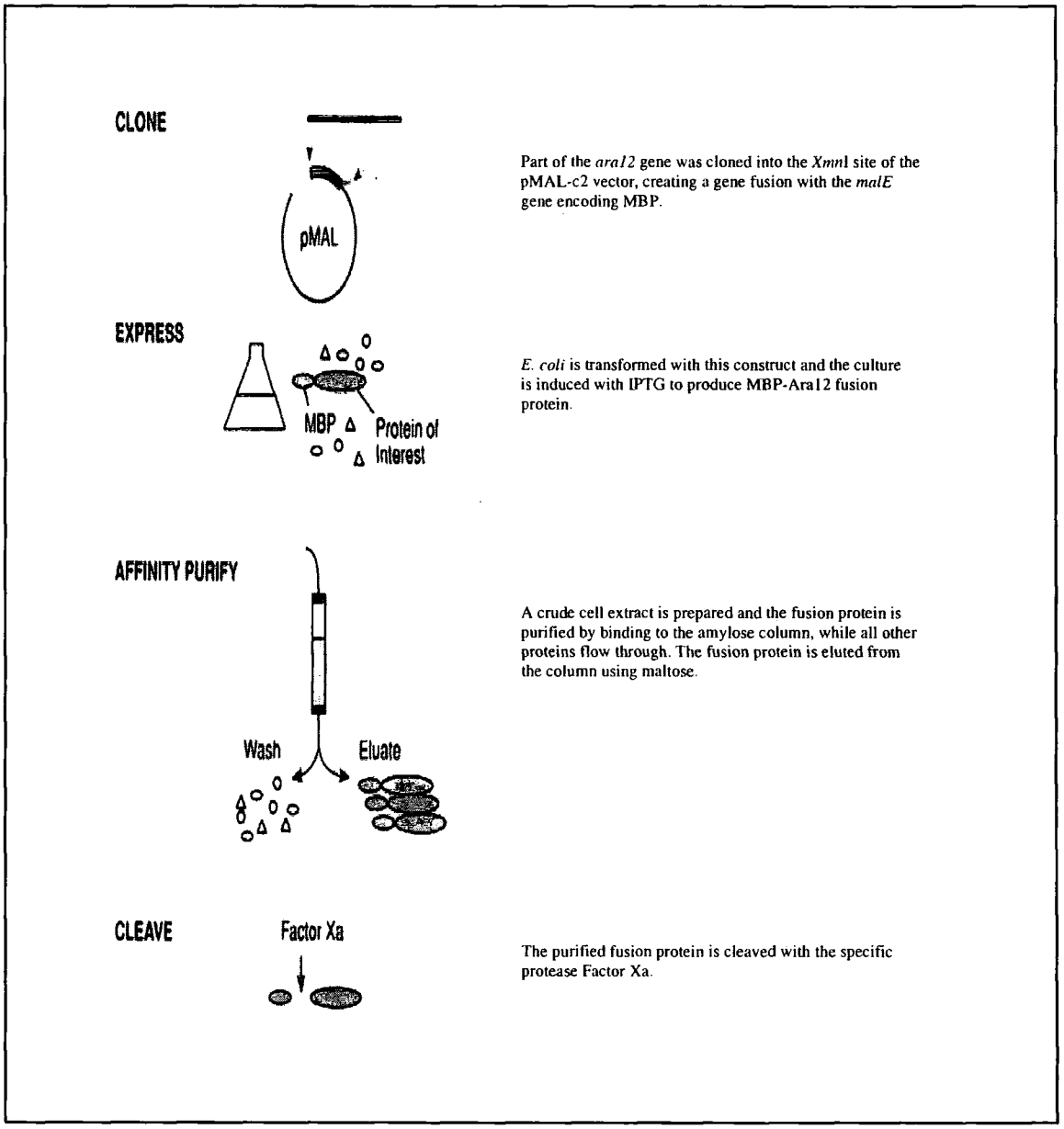


Figure 4.3 Schematic diagram showing the overexpression and purification of part of the Ara12 protease using the pMAL protein fusion and purification system. This diagram was modified from the New England Biolabs Inc. catalogue (1998/99).

cells. The bacterial extract was centrifuged at 12,000 x g at 4°C for 30 minutes to pellet the cell walls, nuclei and other cell debris. The protein concentration of the supernatant was determined by Bradford assay and adjusted to less than 2.5 mg/ml using column buffer. As a guide, 1 g of bacterial cells processed in this way gives about 120 mg of protein. 12 mls of amylose resin (New England Biolabs.) was packed into a column and equilibrated with 20 column volumes of column buffer. The bacterial protein extract was added to the column and washed through with ten column volumes of column buffer. The fusion protein binds to the amylose column via its MBP moiety. The column was washed to remove unbound protein. Bound fusion protein was eluted from the column using two column volumes of 0.5 M maltose. Maltose competes for the binding sites used by MBP to bind to amylose. The eluate was collected and was very cloudy, as it proved to contain high levels of fusion protein.

Proteins in the supernatant from lysed overexpressing cells, the unbound fraction from the amylose column and the eluate from the column were resolved by gel electrophoresis on a 12% SDS polyacrylamide gel (shown in lanes 2-4 of Figure 4.4). The eluate from the column contained large amounts of fusion protein and therefore was diluted 100-fold before loading on the gel. The eluate contained other proteins other than the fusion protein. Therefore the fusion protein was gel purified. Bacterial protein extract was resolved on 10% polyacrylamide gels. Approximately a third of the normal amount of SDS was used in the gels and the running buffer (i.e. 0.03% SDS). Protein bands were visualized using either ChromaPhor stain (purchased from Promega Ltd.) or cupric chloride. A 1/10,000 dilution of ChromaPhor (150 µl for a BioRad mini-gel) was added to the cathode loading

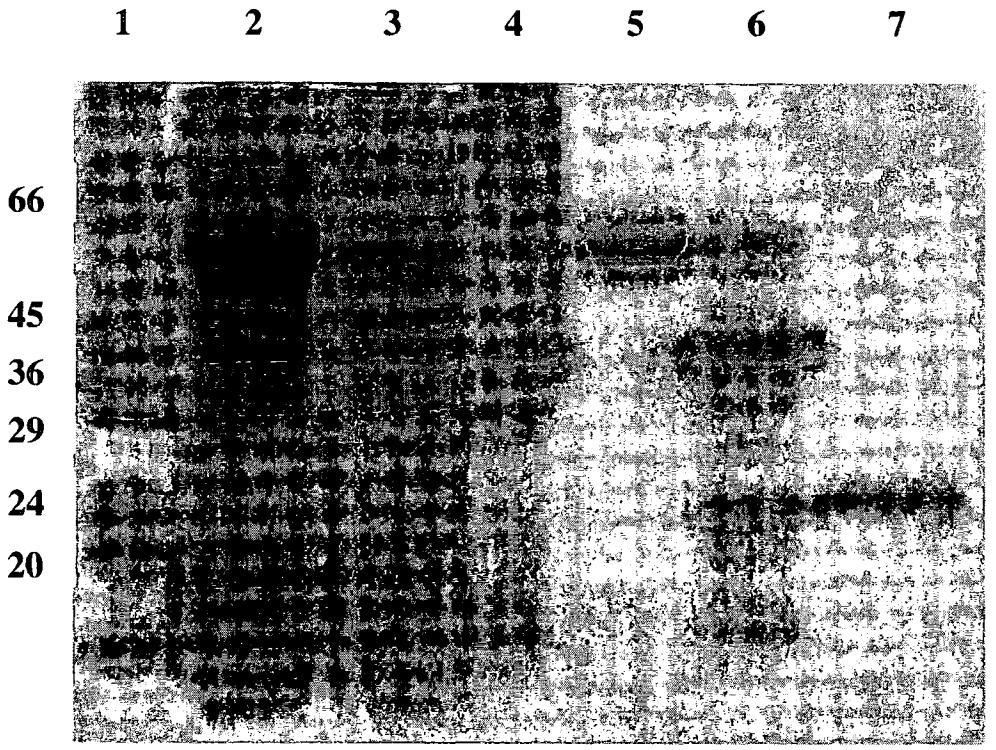


Figure 4.4 Purification of the MBP-Ara12 C-terminal fusion protein by affinity chromatography using amylose resin. Overexpressing XL1 Blue *E. coli* cells were French pressed to lyse the cells. The supernatant proteins were passed through an amylose column, which was washed and eluted with maltose. The fusion protein in the eluate was gel-purified and digested with Factor Xa protease to release the MBP and Ara12 moieties. Lane 1, SDS7 molecular weight markers; lane 2, supernatant from lysed overexpressing XL1 Blue *E. coli* cells; lane 3, unbound fraction from amylose column; lane 4, eluate from amylose column; lane 5, gel-purified fusion protein; lane 6, fusion protein cleaved with Factor Xa protease; lane 7, gel-purified cleavage product (Ara12 C-terminal protein). Molecular weights have been given in kilodaltons.

chamber after loading the sample and the gel was then run as normal. It was found that this method of staining proteins only worked satisfactorily for mini-gels. Larger gels (eg. 20 cm x 20 cm) were stained by placing the gel in 0.3 M cupric chloride for 10-20 min. The gel turns an opaque white colour except where there are protein bands. This negative staining is best viewed against a dark background. The 65 kDa MBP-Ara12 C-terminal fusion protein was viewed on the gel in these ways before excising the band from the gels with a razor blade. Gel slices were cut into small pieces. A Schleicher & Schuell Biotrap electroeluter was used to extract the protein from the polyacrylamide gel. The electroelution was carried out at 60 V overnight at 4°C. Gel purified fusion protein can be seen in lane 5 of Figure 4.4.

Antisera could have been raised to this protein alone, however it was decided to use a combination of the fusion protein and its Ara12 component to generate polyclonal antibodies. This was done to ensure that the resulting antibodies had a high specificity towards the Ara12 protease. Between the MBP and Ara12 portions of the fusion protein there is a Factor Xa protease recognition site. Factor Xa cleaves after this Ile-Glu-Gly-Arg recognition site, separating the two parts of the fusion protein. Purified fusion protein was incubated with Factor Xa in a ratio of 100:1 (w/w) in the following buffer: 20 mM Tris.HCl, pH8.0, 100 mM NaCl, 2 mM $CaCl_2$. All digestions were carried out at 23°C for at least 8 h. Cleavage products were resolved on 12% polyacrylamide gels using 0.03% SDS in the gels and in the running buffer. The 23 kDa Ara12 moiety of the overexpressed fusion protein was excised from these gels with a razor blade, cut into small pieces and recovered from the gel by electroelution as described above. The cleavage products and the gel purified Ara12 C-terminal protein can be seen in lanes 6 and 7 of Figure 4.4.

Protein samples obtained in this way were dialyzed for 4 h against 1x PBS to remove salts and SDS from the solution. The fusion protein and the Ara12 cleavage product was used to raise antibodies which would specifically recognize the Ara12 protease. Masses of these proteins were quantified by densitometry of polyacrylamide gel bands on a BioRad GS-690 imaging densitometer using protein standards of similar length as the samples in question. Binding of Coomassie dye occurs predominantly to aromatic amino acid residues and thus it can be difficult to accurately assess the amount of proteins in this way. In an effort to minimize this effect protein standards of similar molecular weight to those under scrutiny were chosen. Thus the 20 kDa trypsin inhibitor protein was chosen as a standard to determine the amount of the Ara12 C-terminal polypeptide present and BSA was chosen as a standard for the quantification of the full fusion protein. This allowed the protein samples to be assessed for their purity as well as being able to quantify them and very little protein was wasted in this process.

4.3 Generation of polyclonal Ara12 antibodies

A combination of the MBP-Ara12 fusion protein and the purified Ara12 C-terminal protein was used to raise polyclonal antibodies in two rabbits at the Antibody & Assay Technology department of Zeneca Pharmaceuticals at Alderley Edge. Antigens were dialyzed against 0.1xPBS for 18 hours. Rabbits were given subcutaneous injections of either the fusion protein or the Ara12 moiety of the fusion protein in adjuvent as outlined below. I would like to thank Andrew Dinsmore and Jane Bird for organizing the immunizations.

The immunization schedule was as follows:

1) 200 μg fusion protein in complete adjuvent (day 1)

2) 200 μg fusion protein in incomplete adjuvent (day 28)

3) 200 μg fusion protein in incomplete adjuvent (day 56)

4) 40 μg C-terminal portion of Ara12 in incomplete adjuvent (day 84)

5) 40 μg C-terminal portion of Ara12 in incomplete adjuvent (day 140)

The following bleeds were taken from the rabbits during the course of the inoculations:

1) pre-immunization bleed (day 1)

2) first test bleed (day 42)

3) second test bleed (day 70)

4) term bleed (day 98)

5) final bleed (day 154)

On their receipt, the sera were either stored at -80°C or at 4°C after addition of sodium azide to a final concentration of 0.02%. The next section describes the assessment of the quality of the antisera and their use in determining the tissue specificity of the Ara12 protease in *Arabidopsis*.

4.4 Western blot hybridisation analysis

Final bleed serum from rabbits was tested against the same antigens used to inoculate the animals. Purified MBP-Ara12 C-terminal fusion protein (5, 10, 50 and 100 ng) and purified Ara12 C-terminal protein (5, 10 and 50 ng) was resolved on a 12% polyacrylamide gel. The proteins were transferred from the gel to a Hybond-C extra membrane by Western blotting. The blot was probed with a 1/20,000 dilution of final bleed antiserum, followed by a 1/30,000 dilution of anti-rabbit IgG conjugated to horseradish peroxidase (HRP; obtained from Pierce). Immunopositive proteins were detected on the immunoblot as displayed in Figure 4.5. The fusion protein was not detected at the 5 ng loading (see lane 1), but was detected at 10, 50 and 100 ng as a single band of the expected size (see lanes 2-4). The Ara12 C-terminus protein was detected at 5, 10 and 50 ng (see lanes 5-7). In lane 7, two additional immunopositive bands were seen. They correspond in size to the maltose binding protein and the fusion protein. A small quantity of these proteins have most likely become enmeshed in the Ara12 C-terminus protein, even though the latter protein was gel-purified. This may also account for the apparent slight gel shift in molecular weight of the Ara12 moiety of the fusion protein, observed in lanes 6 and 7 of Figure 4.4.

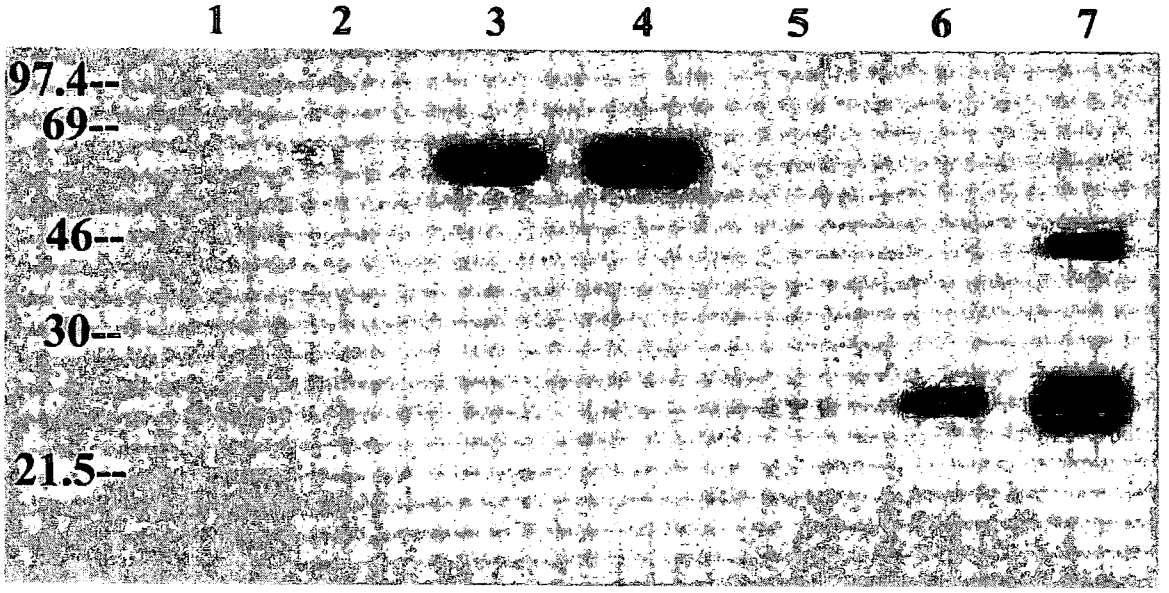


Figure 4.5 Western blot of protein antigens used to raise antibodies to the Ara12 protease. This was probed using the anti-Ara12 antibodies. Lanes 1-4, purified MBP-Ara12 C-terminus fusion protein (5, 10, 50 and 100ng respectively), lanes 5-7, purified Ara12 C-terminus protein (5, 10 and 50ng respectively). The proteins were resolved on a 12% polyacrylamide gel, blotted onto Hybond-C extra and probed first with anti-Ara12 antibodies (1:20,000), followed by anti-rabbit IgG antibodies (1:30,000) conjugated to horseradish peroxidase (HRP). The blot was incubated with SuperSignal chemiluminescent substrate (purchased from Pierce Chemical Company), before exposing the blot to X-ray film. Molecular weights have been given in kilodaltons.

In order to study the location of Ara12 protease in *Arabidopsis* tissues using this antiserum, protein extracts were prepared by grinding fresh tissue in liquid nitrogen with a pestle and mortar. This material was homogenized in an ice cold protein extraction buffer, consisting of 20 mM HEPES, pH7.6, 100 mM potassium acetate. The homogenate was filtered through three layers of Miracloth and used in Western blotting. The protein extracts were resolved on 8% polyacrylamide gels before blotting onto Hybond-C membranes.

Figure 4.6 shows the tissue specificity of Ara12 protease as determined by Western blot hybridisation analysis. In the Western blot used in the results shown in panel A, 13 μg of root, inflorescence stem, immature silique and leaf extract was used in lanes 1-4 respectively. In the blot used in the results shown in panel B, 10 μg of inflorescence stem and immature silique extract was used. Both Western blots were probed in a 1/20,000 dilution of anti-Ara12 serum, followed by a 1/20,000 solution of anti-rabbit IgG-HRP conjugate. The results show that a single immunopositive protein of approximately 80 kDa was detected in *Arabidopsis* tissues. Pre-immune serum used in the way described above on identical blots did not recognize any proteins (data not shown). Hence it was not necessary to immunopurify the anti-Ara12 serum. Immunopositive protein was detected primarily in stem and silique protein extracts. A very faint band of this size was also visible for the leaf extract. This indicates that the actual Ara12 protease is found predominantly in the immature silique and to a lesser extent in inflorescence stems in *Arabidopsis*. It appears that it is also located in leaf tissue at relatively low levels. These results mostly concur with previous published Northern blot hybridisation analysis, which suggested that (judging by *ara12* transcript levels), the Ara12 protease would be likely to be found mainly in silique, flower and rosette leaf tissue, with lower levels found in inflorescence stems, cauline

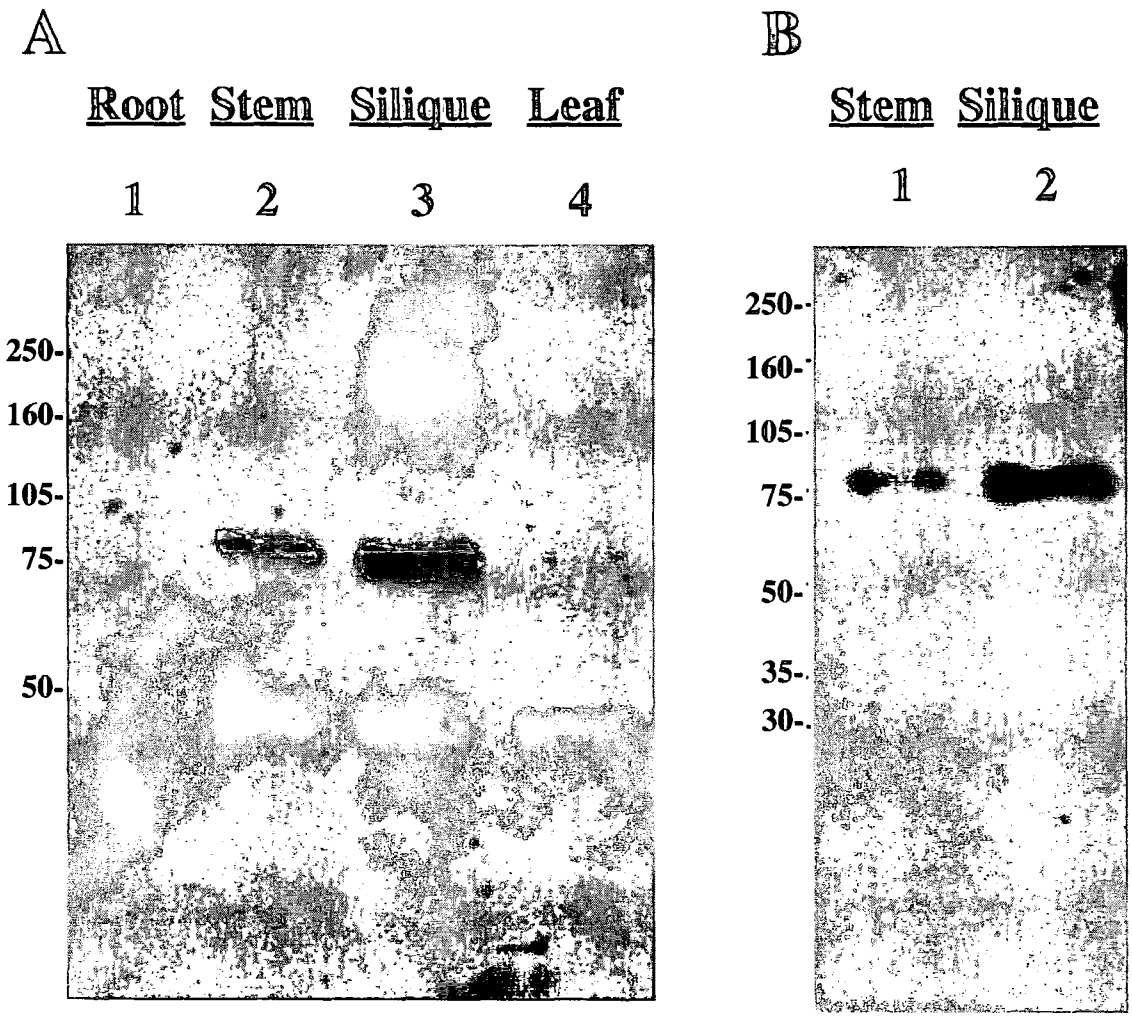


Figure 4.6 Western blot hybridisation analysis showing the tissue specificity of the Ara12 protease in *Arabidopsis thaliana*. Crude protein extracts prepared from the plant tissues indicated were resolved on 10% polyacrylamide gels and blotted onto Hybond-C. The blots were probed with 1:20,000 anti-Ara12 antibody, followed by 1:30,000 anti-rabbit IgG-HRP conjugate. The Western used as shown in panel **A)** contained 13μg of root, stem, silique and leaf protein extract in lanes 1-4 respectively. The Western used as shown in panel **B)** contained 10μg of stem and silique protein extract only in lanes 1 and 2 respectively. Molecular weights have been given in kilodaltons.

leaves and roots (Ribeiro *et al.*, 1995). Unlike this previous work, no distinction was made here between cauline and rosette leaves, when the protein extracts were made, as both were used in leaf extract production. Also flower tissue was not examined here. Nevertheless, both sets of results have shown that the highest levels of Ara12 protease and *ara12* transcripts appeared to be in immature siliques and the lowest levels were found in roots. The results differ in that the anti-Ara12 serum recognizes relatively high amounts of the immunopositive protein in inflorescence stem tissue. Only low levels of Ara12 have been detected in leaf tissue extracts, derived from both cauline and rosette leaves. For the reasons outlined above, the two sets of results are not directly comparable for the leaf samples. These results identified the tissue specificity of this enzyme in *Arabidopsis*, largely confirming a previous examination of *ara12* transcript levels, whilst also demonstrating that the antisera generated were useful in specifically detecting Ara12 protease in complex protein extracts. The presence of *ara12* transcripts does not necessarily mean that the translation products are also present at similar levels in any given tissue. This is the first report of an assessment of the levels of Ara12 protein in *Arabidopsis* tissues.

It was then determined whether recognition of the Ara12 protease in silique extracts by the antiserum could be blocked by addition of MBP-Ara12 fusion protein. This would give a further indication of the specificity of the antiserum towards Ara12 protein. The antibodies should bind to the fusion protein in solution and thus prevent subsequent binding to Ara12 protein immobilized on the Western blot. All but one of the lanes of an 8% polyacrylamide gel were each loaded with 30 μg of immature silique protein extract. The final lane was loaded with molecular weight markers. The proteins were blotted onto a

Hybond-C membrane, after resolving them on the gel. Lanes on the blot were separated from one another using a BioRad Multiscreen clamp, allowing each lane to be incubated with a different solution. Each lane was incubated with a 1/10,000 dilution of anti-Ara12 serum containing 0, 0.1, 0.5, 1, 5, 10, 50, 100 and 200 μg of MBP-Ara12 C-terminus fusion protein, as shown respectively in lanes 1-9 in Figure 4.7. Each of these solutions was made to a total volume of 100 μl. The blot was then washed and incubated with anti-rabbit IgG-HRP conjugate and detection solutions were administered. As can be seen in Figure 4.7, a single band of the same size as that predicted for the Ara12 protease was detected on the immunoblot. As progressively more fusion protein antigen was added to the primary antiserum incubation, the recognition of the immunopositive protein on the blot was masked (see lanes 2-7 in Figure 4.7) until no recognition was detectable on addition of 100 μg of antigen (see lane 8 in Figure 4.7). These results and those displayed in Figure 4.6, demonstrated that the polyclonal anti-Ara12 antiserum was specific for a protein species of a similar size to that predicted for the mature Ara12 protease, which is 67.6 kDa. The immunopositive protein appears to be slightly larger than this and can be accounted for if Ara12 is glycosylated. Extracellular proteins are often glycosylated and several possible glycosylation sites can be identified from its amino acid sequence.

Recognition of the immunopositive protein was reduced by addition of antigen, in a manner which appeared to correlate with the amount of antigen added. These data could be used to determine binding constants of the antisera for the antigen. The results from the experiment shown in Figure 4.7 were used as a criterion on which to judge the quality of the serum produced and influenced a decision to proceed with immunolocalisation by electron microscopy.

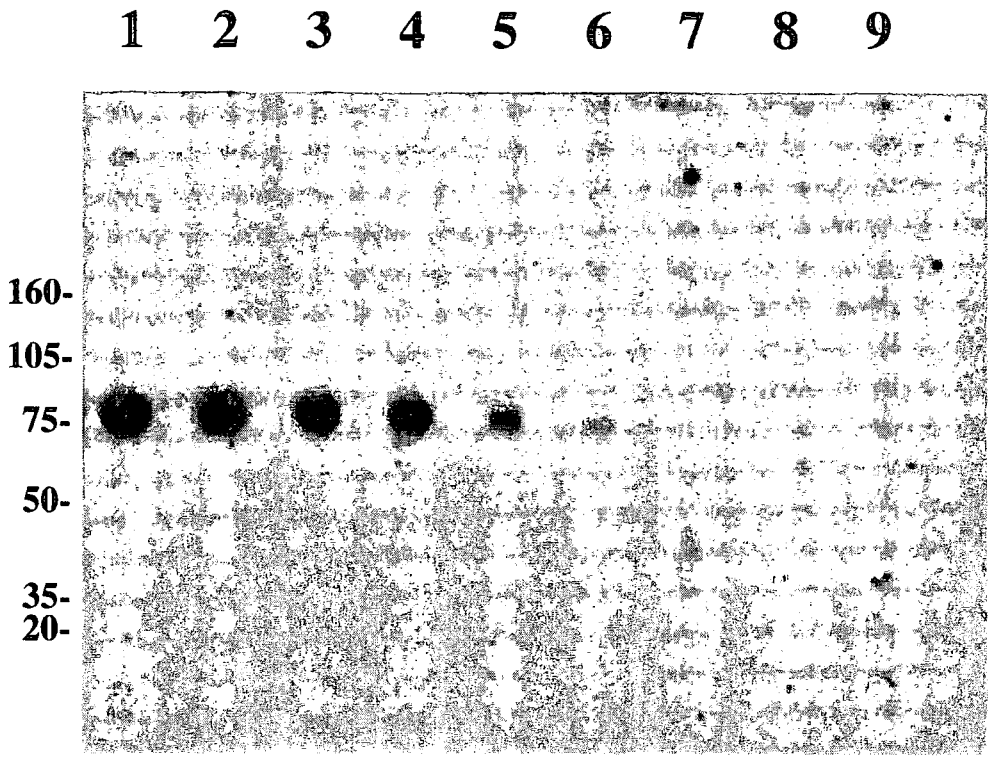


Figure 4.7 Western blot of silique protein extract masked with progressively more antigen. Each lane of the blot contained 30μg of crude silique protein extract which was resolved through an 8% polyacrylamide gel. The blot was clamped using a multiscreen to separate the lanes. Each lane was incubated with anti-Ara12 antiserum (1:10,000) mixed with varying amounts of MBP-Ara12 C-terminus fusion protein, followed by anti-rabbit IgG antiserum (1:20,000) conjugated to HRP. Lanes 1-9, the amount of MBP-Ara12 C-terminus fusion protein added to the primary antibody was 0, 0.1, 0.5, 1, 5, 10, 50, 100 and 200μg respectively. The blot was then incubated with SuperSignal chemiluminescent substrate, before exposing the blot to X-ray film. Molecular weights have been given in kilodaltons.

4.5 Immunolocalisation by electron microscopy

In order to look more closely at the cellular and subcellular localisation of the enzyme, immunochemical methods combined with the resolution of electron microscopy was used. Several *Arabidopsis* plant tissues, such as young and mature siliques and young inflorescences were fixed and embedded in LR White resin. The samples were sectioned and incubated with either anti-Ara12 serum (1/100) or preimmune serum (1/100) and then anti-rabbit IgG conjugated to 20 nm gold particles (1/100). I would like to thank Stefan Hyman at the University of Leicester for sectioning and staining samples and recording the electron microscope images presented here.

The localisation of Ara12 protease was investigated using longitudinal and transverse sections of *Arabidopsis* tissues. Transverse sections through an *Arabidopsis* inflorescence are shown in Figures 4.8 – 4.10. Figure 4.8 shows an electron micrograph of part of a transverse section of an *Arabidopsis* inflorescence incubated with preimmune serum and anti-rabbit IgG conjugated to gold particles. Gold particles were only rarely found to be associated with the cell wall, intercellular spaces, the cytoplasm or cellular organelles, such as chloroplasts and mitochondria. Figures 4.9 and 4.10 show electron micrographs of parts of transverse sections of an *Arabidopsis* inflorescence incubated with anti-Ara12 serum and anti-rabbit IgG conjugated to gold particles. Gold particles were found in the intercellular spaces, but were only rarely found to be associated with the cell wall, the cytoplasm or cellular organelles. This strongly suggests that Ara12 protease is found in the intercellular spaces in the inflorescences of *Arabidopsis.*

There is some evidence that Ara12 is also found in intercellular spaces in siliques (data not shown), however further experimentation would be required to confirm this. Gold

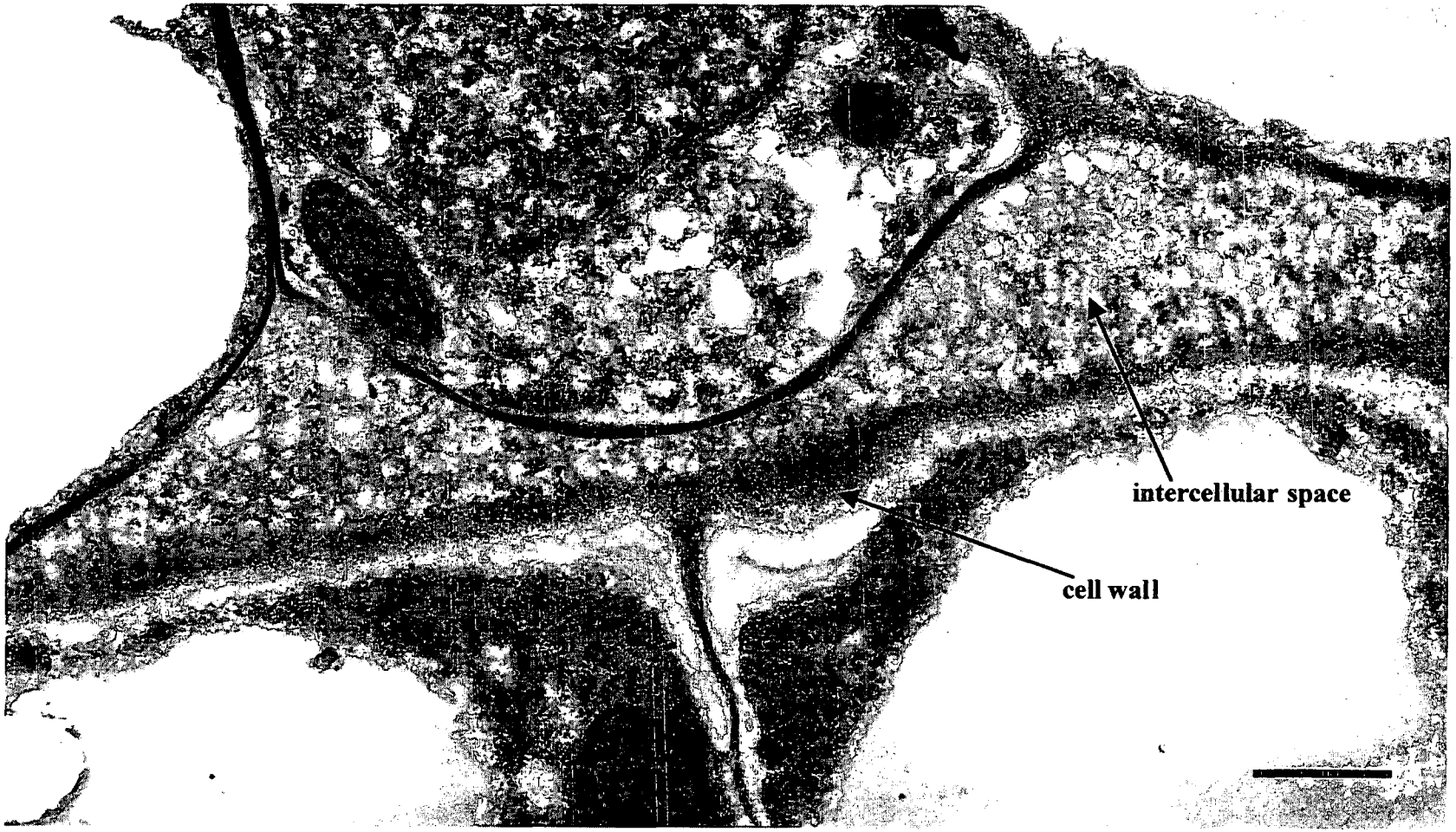


Figure 4.8 Electron micrograph of part of a transverse section of an inflorescence from *Arabidopsis* (23,500X magnification) incubated with preimmune serum (1/100). The section was subsequently incubated with anti-rabbit IgG conjugated to gold particles (20nm in diameter). Gold particles were found only very rarely in intercellular spaces, the cell wall, cytoplasm and in cellular organelles, such as chloroplasts and mitochondria. Scale bar = 850nm.

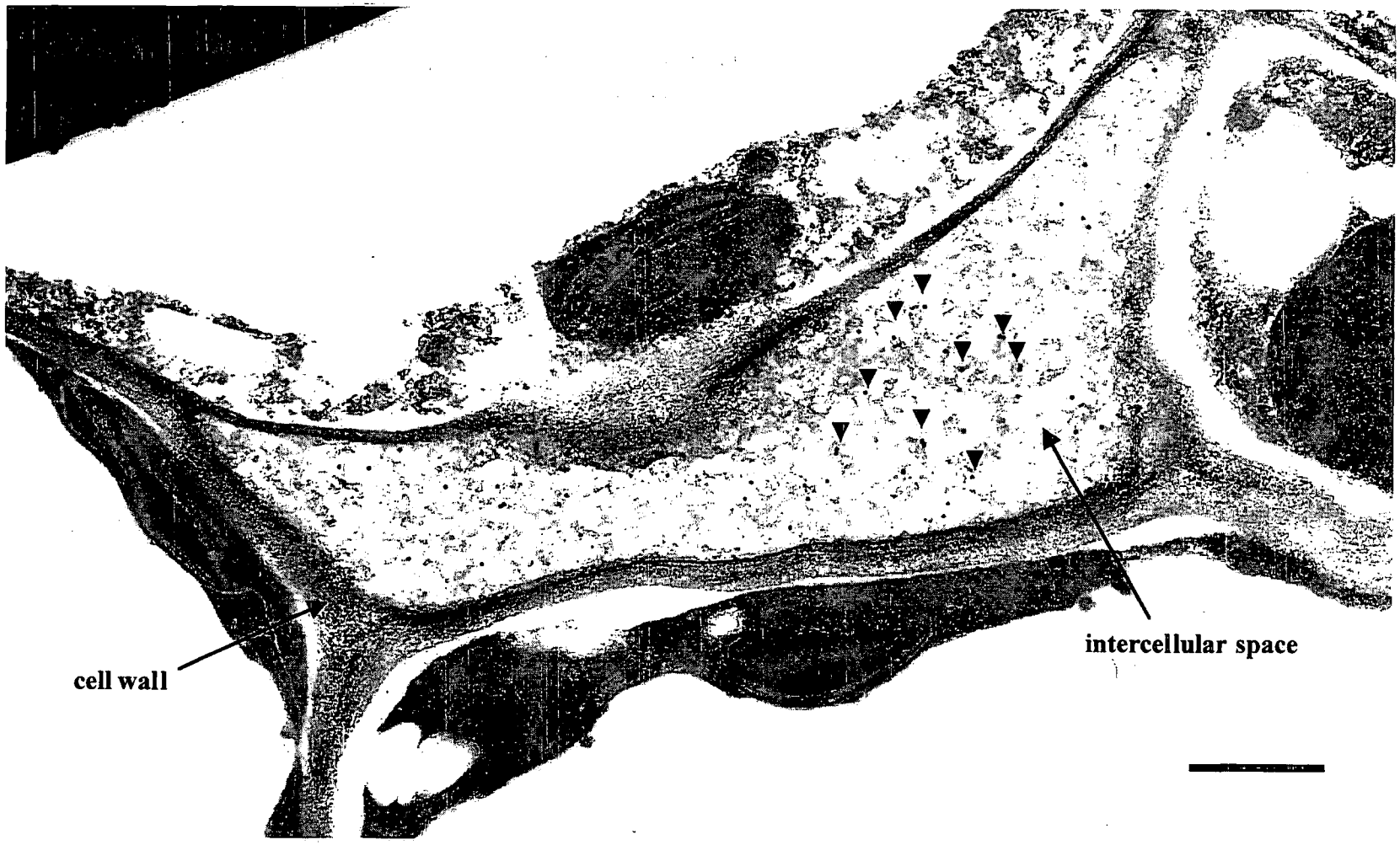


Figure 4.9 Electron micrograph of part of a transverse section of an inflorescence from *Arabidopsis* (23,500X magnification) incubated with anti-Ara12 serum (1/100). The section was subsequently incubated with anti-rabbit IgG conjugated to gold particles (20nm in diameter). Gold particles were found in the intercellular spaces (indicated above by arrowheads), and only very rarely in the cell wall, cytoplasm and in cellular organelles, such as chloroplasts and mitochondria. Scale bar = 850nm.

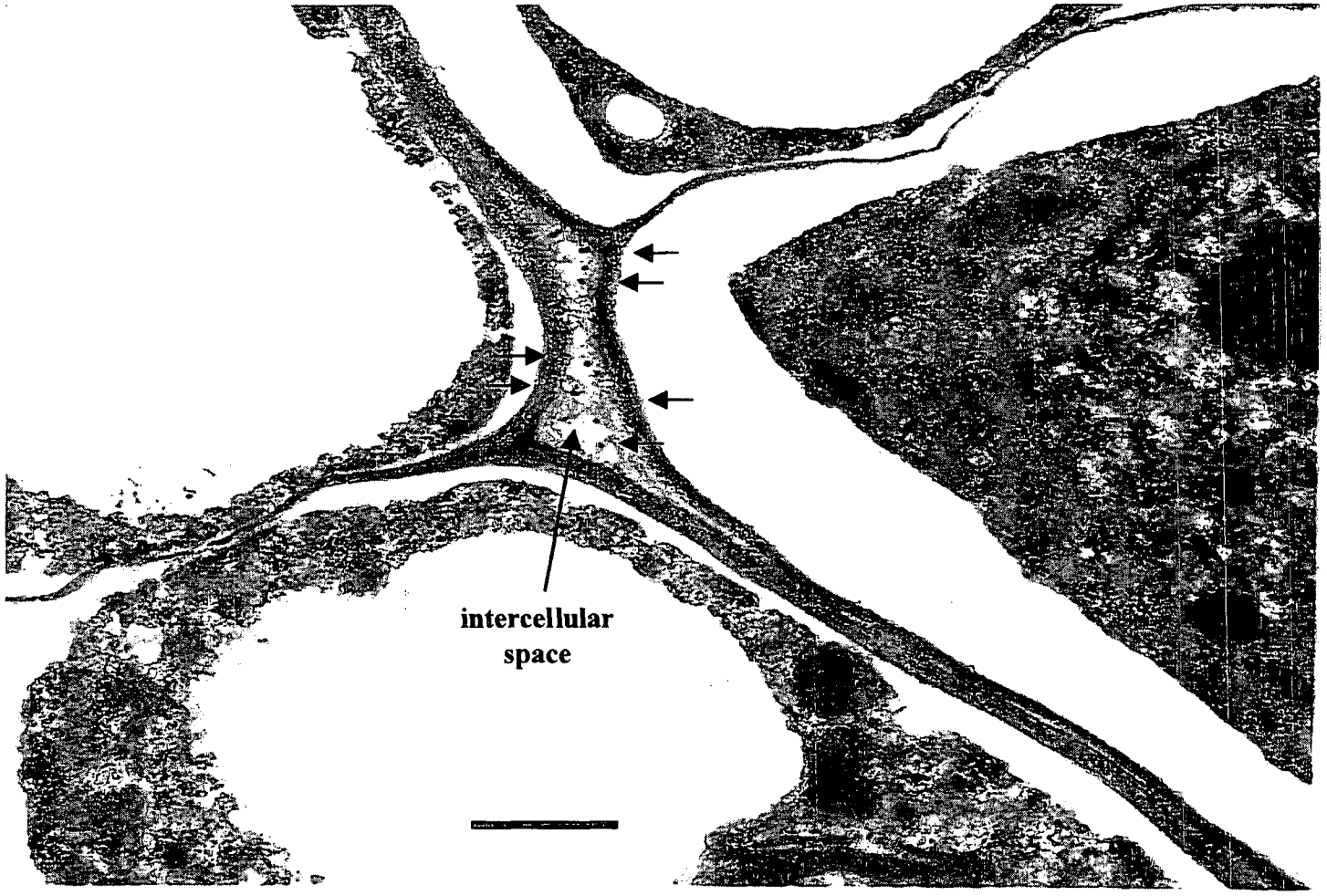


Figure 4.10 Electron micrograph of part of a transverse section of an inflorescence from *Arabidopsis* (23,500X magnification) incubated with anti-Ara12 serum (1/100). The section was subsequently incubated with anti-rabbit IgG conjugated to gold particles (20nm in diameter). Gold particles were found in the intercellular spaces (indicated above by arrows), and only very rarely in the cell wall, cytoplasm and in cellular organelles, such as chloroplasts and mitochondria. Scale bar = 850nm.

particles have also been found in the cell wall of cells found in longitudinal sections of siliques, which had been incubated with anti-Ara12 serum, however some gold particles were also found in the cell wall of control sections incubated with preimmune serum (data not shown). It has not been possible to determine whether Ara12 is found in the cell wall, using this approach, because antibodies in the preimmune serum can associate with the cell wall.

4.6 Conclusions

In the main body of this chapter of results, some difficulties encountered in overexpressing the complete *ara12* gene product were discussed. Evidence exists to suggest that Ara12 protease displays activity to a broad range of protein substrates. Therefore cell death of bacteria expressing the enzyme should not perhaps be a surprising outcome.

Polyclonal antibodies have been raised against a mixture of the C-terminal 216 amino acid residues of the Ara12 protease and a fusion protein consisting of this protein fused to maltose binding protein. The antisera generated have been shown to be highly specific for the Ara12/Slpa protease as judged by their recognition of a single protein band of the expected size in *Arabidopsis* tissue extracts. Furthermore, the recognition could be blocked as progressively more antigen (purified fusion protein) was added to the primary antibody wash, as more and more of the antigen in solution competed for the protein immobilised on the Western blot.

Ara12/Slpa protease was found predominantly in silique tissue and to a lesser extent in stem tissue. Leaf and root tissue appeared to contain very little of the protease. This tissue specificity found using the antisera generated largely matched previously reported Northern blot analysis results (Ribeiro *et al.*, 1995). Western blot analysis, however implied that inflorescence stems contained higher levels of Ara12 than was suggested by this previous study.

Immunolocalisation studies were conducted by electron microscopy using gold labelling. This technique was employed to examine sections of silique and stem tissue for the presence of Ara12. The Ara12 protease appears to be located in the intercellular spaces in the inflorescences of *Arabidopsis*. This is in accordance with the predicted targeting of the Ara12 protein, as the presence of a putative N-terminal signal peptide suggests that the mature protease has an extracellular location. Further work is required to establish whether Ara12 is found in intercellular spaces in silique tissue. It has not been possible to determine whether Ara12 is also found in the cell wall, as antibodies in the preimmune serum can associate with the cell wall.

Chapter 5

Purification of Ara12/Slpa subtilisin-like protease from *Arabidopsis* cell suspension culture filtrate

5.1 Introduction

Although much is known about prokaryotic subtilisins, information regarding plant subtilisin-like proteases in general is relatively scarce. There are several substantial differences between these plant serine proteases and typical subtilisins, such as subtilisin BPN' or subtilisin Carlsberg (Bogacheva, 1999). Characteristically the plant enzymes have large molecular masses with long C-terminal regions and contain several cysteine residues, unlike their bacterial counterparts. A small number of subtilisin-like proteases have been purified including cucumisin from melon fruit (Kaneda and Tominaga, 1975) and taraxalisin from dandelion root (Rudenskaya *et al.*, 1998), however no subtilisin-like protease has been purified from *Arabidopsis*, an important model plant. There is growing evidence for the presence of a large number of subtilisin-like proteases in *Arabidopsis* plants (see Table 3.1). Purification of Ara12 protease would enable the first biochemical study of a member of this group of enzymes.

Serine proteases are increasingly being purified from plant cells, for example a serine protease has recently been purified from suspension cultured soybean cells (Guo *et al.*, 1998). A previous investigation into plant extracellular proteins was conducted at the University of Durham in which a subtilisin-like protease was identified from *Arabidopsis* (Robertson *et al.*, 1997). This study compared primary cell wall proteins from suspension cultured cells of *Arabidopsis*, carrot, French bean, tomato and tobacco. The proteins were extracted by washing whole cells successively in aqueous $CaCl_2$, CDTA, DTT, NaCl and borate solutions. In addition the filtrate from growing suspension cultured cells was examined. Protein profiles from the samples were examined on 10% polyacrylamide gels. Twelve main culture filtrate proteins (between approximately 20-150 kDa) were selected

for sequencing and eleven N-terminal sequences were obtained. One of the proteins found in *Arabidopsis* culture filtrate (ACF) was predicted to be a subtilisin-like protease (Ara12), and has been predicted from the near full-length *ara12* cDNA that has been isolated by PCR from an *Arabidopsis* cDNA library (Ribeiro *et al.*, 1995). The amino acid terminus of the sequenced ACF protein corresponds to the predicted cleavage site at which the Ara12 proenzyme is processed to its mature form (Siezen *et al.*, 1991). This proved that the mature Ara12 protease is found in the filtrate of *Arabidopsis* cell cultures, as a moderately dominant protein (when compared to the relative quantities of the other proteins present in the sample). The aim of this chapter was to purify functional Ara12 subtilisin-like protease from *Arabidopsis* in order to study aspects of the biochemistry of this extracellular enzyme. Using the same cell cultures, protease activity was detected in concentrated culture filtrate (as will be shown in the next section), and this was examined as a potential starting material for the purification.

5.2 Detection of protease activity in *Arabidopsis* cell culture filtrate

As outlined above, Ara12 protease was known to be found in *Arabidopsis* cell culture filtrate. The protease activity of the culture filtrate was determined using assays based on azocasein or casein labelled with fluoroisothiocyanate (FITC). The assays were conducted with a minor modification of a previously published method (Twining, 1984; see Chapter 2). No protease activity could be detected in *Arabidopsis* culture filtrate with the more sensitive FTC-casein assay, even after a 16 hour incubation at 37°C (see Figure 5.1).

In order to detect the protease activity *Arabidopsis* culture filtrate was concentrated by ultrafiltration at 4°C using an Amicon 8200 stirred filtration cell (Amicon Corporation). An ultrafiltration membrane with a molecular weight cut-off of 10 kDa was used to retain the larger proteins. Culture filtrate was filtered through the membrane under pressure using nitrogen gas and the samples were concentrated. In this way 20-fold and 60-fold (by volume) culture filtrate samples were generated. The relative protease activities of these samples was compared to non-concentrated culture filtrate using FTC-casein. One unit (1 U) of FTC-casein degrading activity is defined as the amount of enzyme required to produce 1.0 unit of Ex485 Em538 fluorescence increase under the standard conditions of the assay. These units were adopted throughout this work and have been used by others during the purification of tomato P-69A protease (Vera and Conjero, 1988). Protease activity was detected in 20-fold and in 60-fold ACF after a 16 hour incubation at 37°C (see Figure 5.1). As the Ara12 protease had been shown to be present in ACF, the next question to arise was whether the protease was partially or wholly responsible for this activity. This question was addressed by immunoblot analysis with anti-Ara12 serum generated previously in this study. Figure 5.6 shows that Ara12 can be detected in 60x ACF, while re-

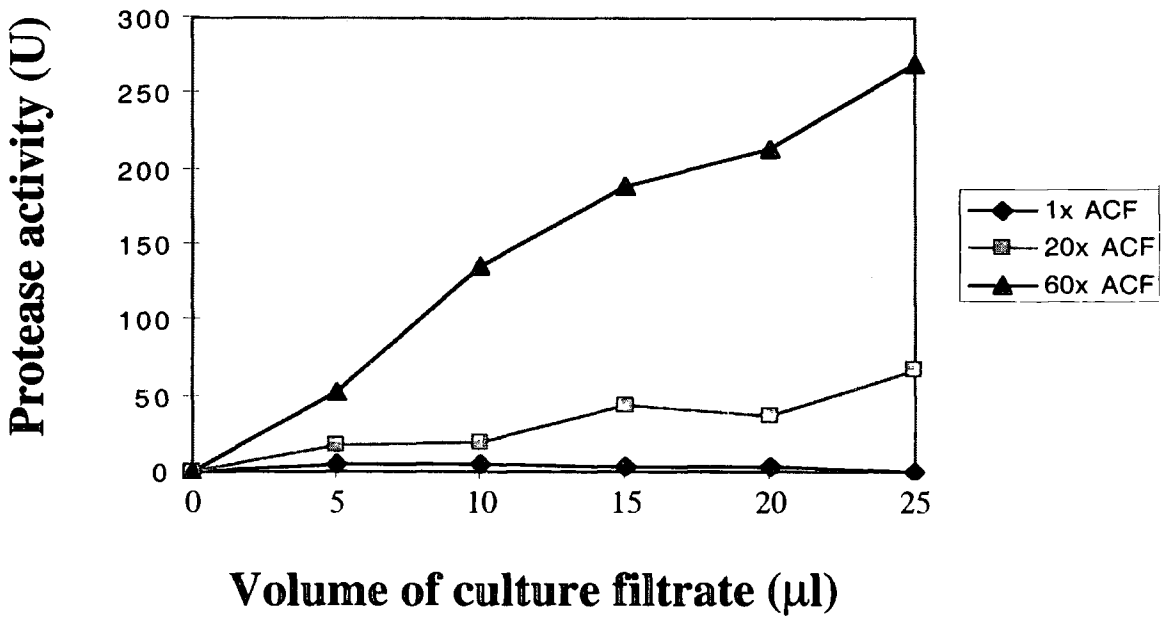


Figure 5.1 Detection of protease activity in *Arabidopsis* cell suspension culture filtrate concentrated by ultrafiltration. Six days after subculture, a cell suspension culture was filtered through two layers of Miracloth. The culture filtrate was concentrated 20-fold and 60-fold by volume by ultrafiltration in a stirred cell using a 10kDa molecular weight cut-off membrane at 4°C. Protease activities of different volumes of unconcentrated *Arabidopsis* culture filtrate (1x ACF) and 20-fold and 60-fold concentrated *Arabidopsis* culture ultrafiltrate (20x ACF and 60x ACF) were measured in a 50µl protease assay using FTC-casein.

maining undetectable in 1x ACF. This underlines the protease assay findings and implies that an activity associated with the Ara12 protease can be located by Western blotting. The next stage was to see if the Ara12 protease activity could be separated from the bulk of the other proteins in the solution. This was initially done using concentrated ACF as a starting material before repeating the process on a larger scale with 1x ACF by a batch binding process. The first task was to find a gel matrix suitable for this purpose, as is described in the next sections.

5.3 Anion exchange chromatography of *Arabidopsis* cell culture ultrafiltrate

After having detected protease activity in *Arabidopsis* culture ultrafiltrate, the next step taken was to see whether an initial ion exchange chromatography step would be useful in the purification of Ara12/Slpa protease. Ion exchange chromatography (IEC) utilizes principally the electrostatic interaction between the different net charges found on proteins and ion exchange adsorbents, which can be positively (anion exchangers) or negatively (cation exchangers) charged (Himmelhoch, 1971). The interactions decrease as the ionic strength of the buffer is increased, resulting in desorption of protein bound to IEC gel matrices. It is a much used and powerful technique, as proteins with only minor differences in charge can be separated efficiently.

Six days after subculture, *Arabidopsis* suspension cell cultures were filtered through Miracloth and filter paper. The filtrate from these cultures was concentrated using an Amicon 8200 stirred cell at 4°C. A PM10 membrane was used in the microconcentrator, with a 10 kDa molecular weight cut off, so that proteins of this molecular weight or larger were retained. Culture filtrate was concentrated approximately 60 times by volume, and was then dialyzed overnight at 4°C against 20 mM Tris.HCl, pH7.5. After centrifuging at 40,000 x g for 10 minutes at 4°C, the pH and conductivity of the supernatant was determined (pH7.5 and 0.7 mS respectively). Half a millilitre of supernatant was loaded onto a 1.6 mm/5 cm Mono-Q anion exchange column, which had been equilibrated with 4 column volumes of 20 mM Tris.HCl, pH7.5. A flow rate of 100 µl/min was maintained. After loading the sample, the column was washed with 4 column volumes of 20 mM Tris.HCl, pH7.5 and the unbound protein fraction was collected. Bound proteins were

eluted with a linear gradient of sodium chloride (0-0.4 M) over 15 column volumes, followed by 2 column volumes of 1 M NaCl in 20 mM Tris.HCl, pH7.5. Twenty 100 µl fractions were collected. The load, the column and the fractions were kept at 4°C to minimize proteolytic degradation during this procedure.

A chromatogram of this fractionation is shown in Figure 5.2. The anion exchange chromatography procedure outlined here was repeated having raised the pH of the load and the column and elution buffers to pH8.5. A chromatogram of this second fractionation is shown in Figure 5.4. Fractions from these chromatographic steps were assayed using the FTC-casein assay described. Fractions displaying high levels of protease activity were run on 8% polyacrylamide gels, which were silver stained to detect the small quantities of proteins present. Figure 5.3 shows the SDS PAGE analysis of active fractions from the Mono-Q HPLC fractionation carried out at pH7.5. Figure 5.5 shows SDS PAGE analysis of active fractions from both these Mono-Q fractionations performed (i.e. at pH7.5 and pH8.5), for comparison.

Protein fractions were resolved on a separate 8% polyacrylamide gel and blotted onto Hybond-C before probing with anti-Ara12 serum (see Figure 5.6). This demonstrated that the fractions showing protease activity against FTC-casein specifically contained the Ara12 protease.

As will be seen later in this chapter, the main proteolytic activity in the *Arabidopsis* suspension culture filtrate studied was caused by the action of Ara12 protein, however a second proteolytic activity was identified during the course of the purification of Ara12 protease. The source of this activity was not determined. Although it is quite likely that other protease activities would be found in the culture filtrate, two peaks of activity stood

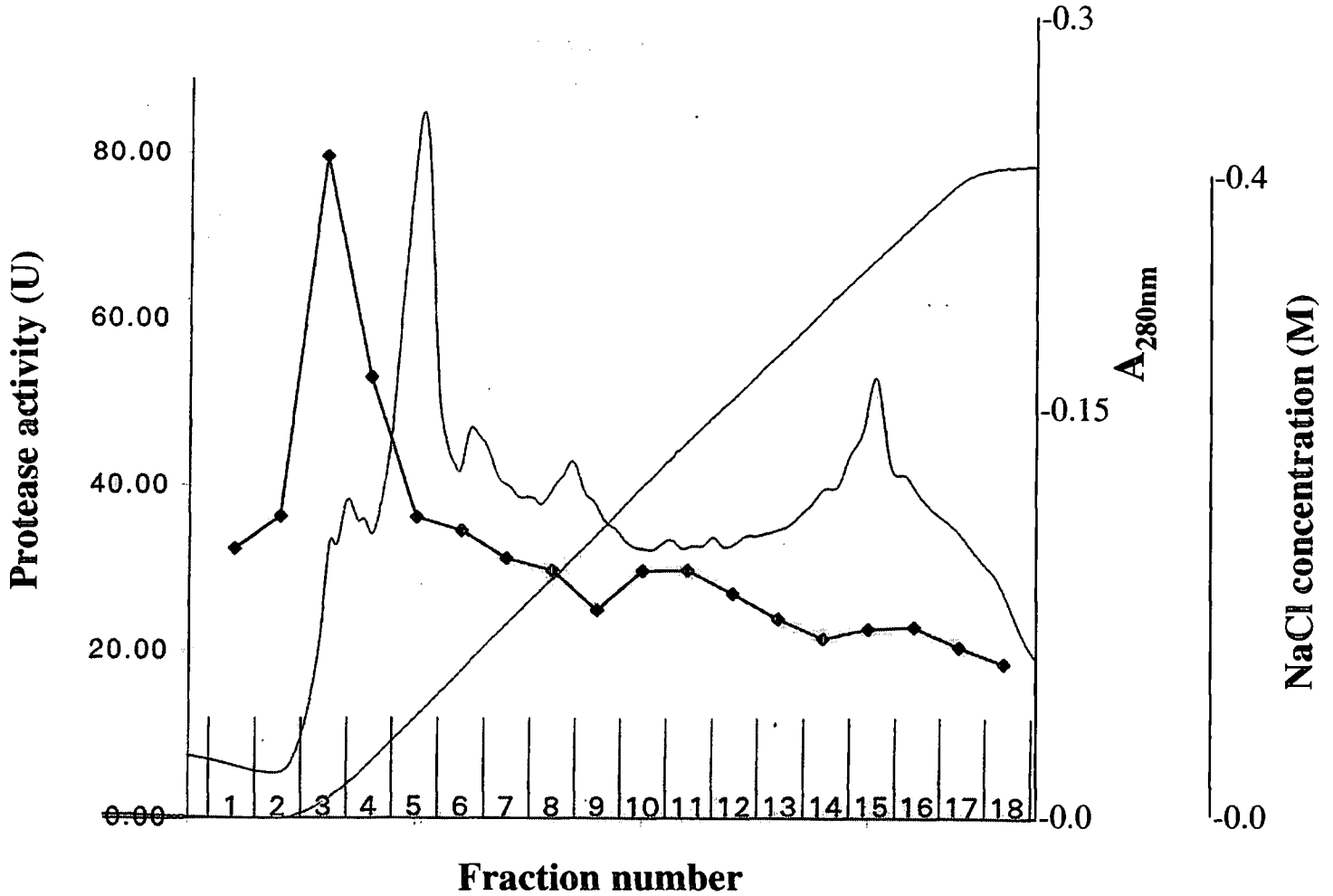


Figure 5.2 Chromatogram of *Arabidopsis* culture ultrafiltrate fractionated by HPLC on a Mono-Q column at pH7.5. 60x ACF was dialyzed against 20mM Tris.HCl, pH7.5 and loaded onto a 1.6/5 Mono-Q HPLC column, which had been equilibrated with the same buffer. Proteins were eluted on a linear gradient of 0-0.4M NaCl as shown above. The protease activity of the 100µl fractions was determined using FTC-casein and is shown by the line with the black diamonds. The other line on the chromatogram shows the absorbance at 280nm (A_{280nm}).

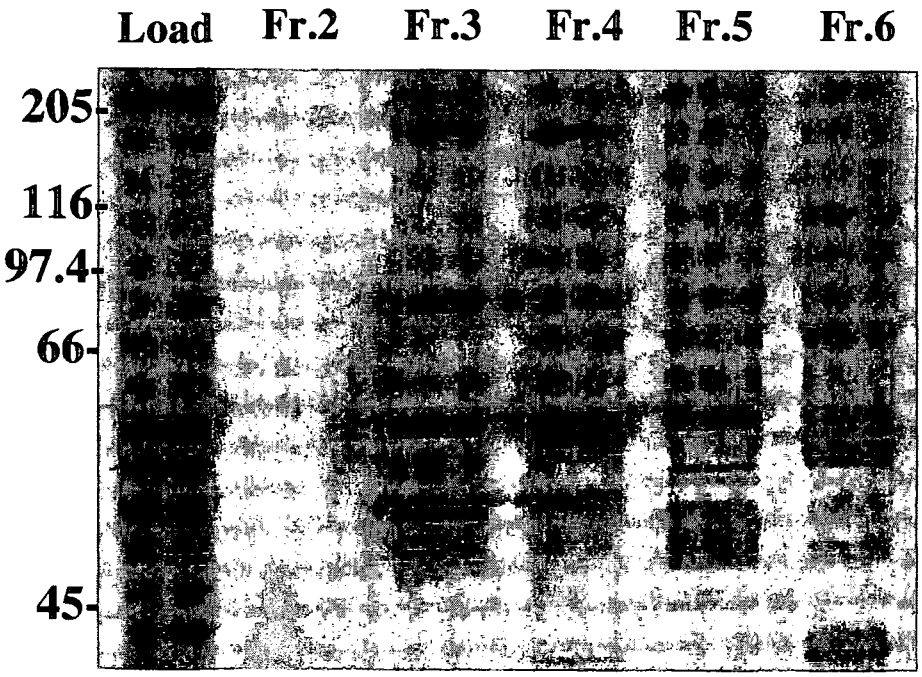


Figure 5.3 SDS-PAGE analysis of the Mono-Q fractionation performed at pH7.5. 10µl of the load and Mono-Q fractions 2-6 were resolved on an 8% polyacrylamide gel. Proteins were visualized by silver staining. The molecular masses are given in kDa. Fraction 3 showed the highest level of protease activity of the fractions produced.

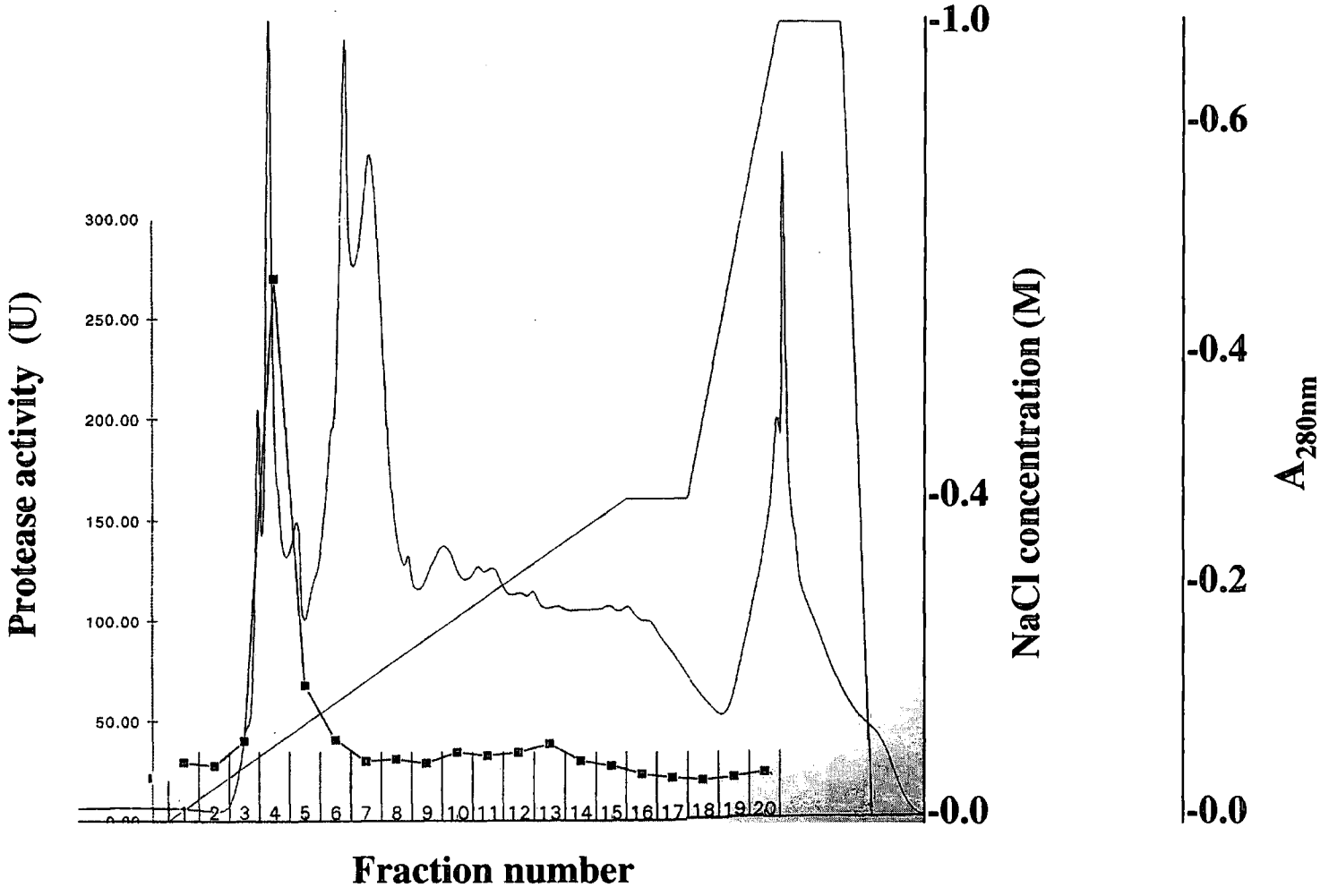


Figure 5.4 Chromatogram of *Arabidopsis* culture ultrafiltrate fractionated by HPLC on a Mono-Q column at pH8.5. 60x ACF was dialyzed against 20mM Tris.HCl, pH8.5 and loaded onto a 1.6/5 Mono-Q HPLC column, which had been equilibrated with the same buffer. Proteins were eluted on a linear gradient of 0-0.4M NaCl as shown above. The protease activity of the 100µl fractions was determined using FTC-casein and is shown by the line with the black diamonds. The other line on the chromatogram shows the absorbance at 280nm (A_{280nm}).

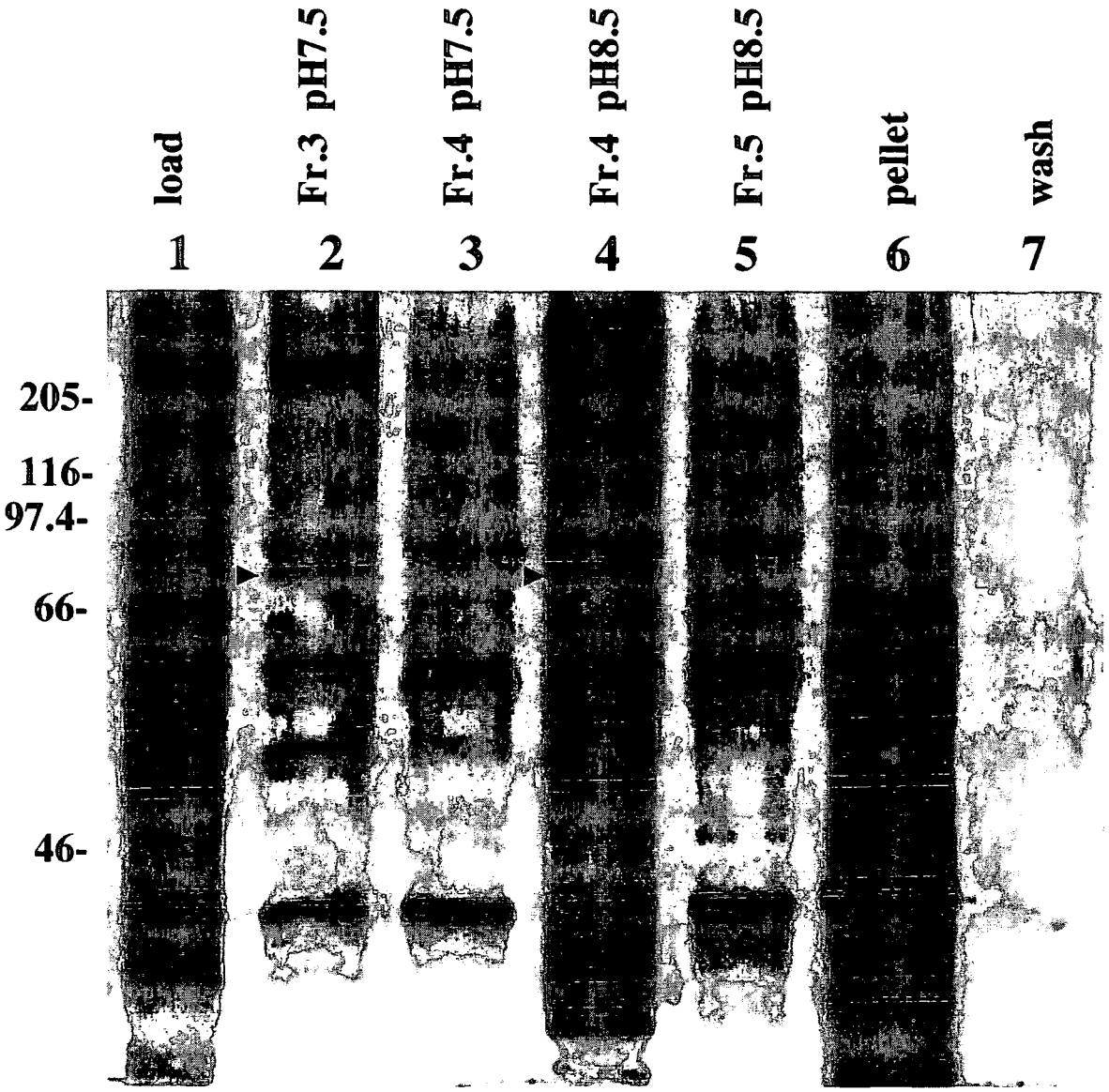


Figure 5.5 SDS-PAGE analysis of Mono-Q fractions showing protease activity. The load, wash and fractions showing high levels of protease activity from the Mono-Q fractionation performed at pH7.5 and pH8.5 were loaded onto a 10% polyacrylamide gel. A 60x ACF sample prepared from *Arabidopsis* cell culture filtrate by ultrafiltration was centrifuged at 40,000 x g for 10 minutes. The pellet was resuspended in 20mM Tris.HCl, pH7.5 and was also examined by SDS-PAGE, whilst the supernatant was used as the load in both fractionations. Lane 1, 10µl load; lane 2-3, 10µl fractions 3 and 4 of Mono-Q fractionation at pH7.5; lanes 4-5, 10µl fractions 4 and 5 of Mono-Q fractionation at pH8.5; lane 6, pellet from centrifuged 60x ACF; lane 7, 10µl combined wash from both fractionations. Resolved proteins were visualized by silver staining. Molecular masses are given in kDa. Fraction 3 and 4 of the fractionations performed at pH7.5 and pH8.5, respectively, showed the highest level of protease activities. The mature Ara12 protease is indicated with arrowheads.

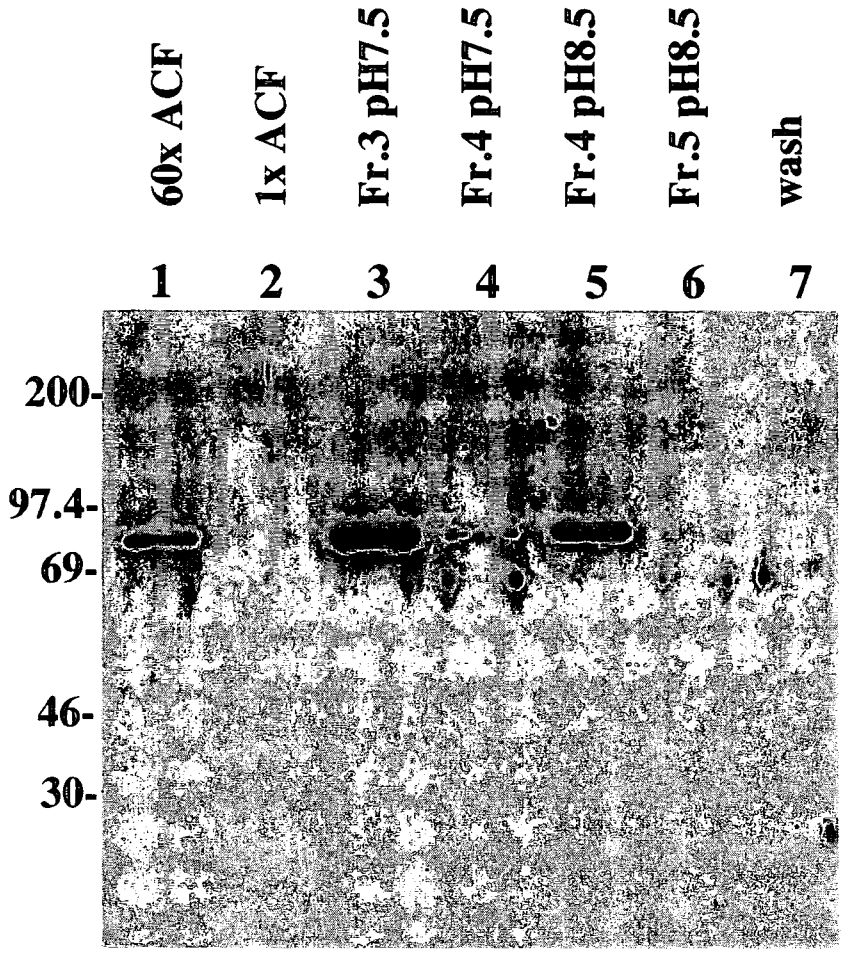


Figure 5.6 Western blot hybridisation analysis of *Arabidopsis* culture filtrate and ultrafiltrate and Mono-Q active fractions. 60x ACF ultrafiltrate was used as the load for the Mono-Q fractionations. Lane 1, 60x ACF; lane 2, 1x ACF; lanes 3-4, fractions 3-4 of Mono-Q fractionation at pH7.5; lanes 5-6, fractions 4-5 of Mono-Q fractionation at pH8.5; lane 7, combined wash from both Mono-Q runs. 10µl of each sample was resolved on a 10% polyacrylamide gel, blotted onto Hybond-C extra and probed with a 1:20,000 dilution of anti-Ara12 serum, followed by a 1:50,000 dilution of goat anti-rabbit IgG-HRP conjugate.

out clearly.

Although the intensity of the silver stained bands is not precisely quantifiable, an indication is given that the fractions with the highest protease activities contain the greatest amounts of the protein of the expected size (approximately 80 kDa). Immunoblot analysis showed that samples from both fractionations which showed the highest activities also proved to contain an immunopositive protein of the same molecular weight as that expected for Ara12. The conclusion drawn from this was, firstly, that *Arabidopsis* cell culture filtrate did indeed contain a proteolytic activity associated with the subtilisin-like protease Ara12/Slpa. This activity proved to be the dominant proteolytic activity caused by enzymes in the filtrate as shown by fractionation of filtrate proteins on Mono-Q anion exchange resin. Fractions displaying relatively high protease activities proved to have relatively large amounts of a protein with a molecular weight similar to that predicted for Ara12. When probed with antibodies believed to be extremely specific for the Ara12 protein, these fractions also contained an immunopositive protein of the expected molecular weight.

5.4 Batch binding using Q Sepharose FF anion exchange resin

Having established that an ion exchange resin based on the quaternary ammonium ion (Q) could be used to good effect to isolate Ara12 from concentrated ACF, it was speculated that it should be possible to achieve purification using such a resin with *Arabidopsis* culture filtrate, which had not been concentrated. Apart from being time-consuming, when dealing with large volumes of ACF, the concentration step could result in a loss of yield, due to autolytic activity and adherence of the protease to the ultrafiltration membrane. Effectively the concentration step could be performed rapidly and efficiently on large quantities of ACF using an anion exchange resin. This was done by adding equilibrated resin directly to ACF which had been diluted with deionised water. ACF was diluted to lower the ionic strength of the solution. If the ACF had been at too high an ionic strength, then the proteins would not be able to bind to the resin. In this case the proteins would of course remain in the solution and would not be extracted. The culture filtrate proteins could then be eluted in a small volume. In this way Ara12 could potentially be purified from a readily available source with no need for an ultrafiltration step.

Initially a small scale batch binding experiment was conducted to confirm the optimum pH for the extraction of the protease activity from 60x ACF with Q Sepharose FF. This resin was equilibrated and made up to a 50% slurry with solutions of 10 mM Tris.HCl of varying pH (6.0, 6.5, 7.0, 7.5, 8.0, 8.5 and 9.0). 200 µl of this slurry was added to 150 µl of 60x ACF which had been dialyzed against 10 mM Tris.HCl of the corresponding pH. Tubes were left shaking gently at room temperature for 40 minutes, before centrifuging the samples at 13,000 x g for 2 minutes and assaying the supernatant for protease activity

using an FTC-casein assay. The results, shown in Figure 5.7, suggested that over the pH range tested most protease activity was removed at pH8.5 (with 80% removed compared to a gcontrol which was not extracted with Q Sepharose).

Arabidopsis cell cultures were filtered six days after subculture, having grown at 24°C, shaking at 130 rpm in a light regime of 16 hours of light and 8 hours of darkness. Fresh filtrate was diluted 4-fold with ice cold deionised water to give a final volume of 1360 mls. The conductivity of the solution dropped from 4.1 mS/cm to 1.2 mS/cm as a result. The pH of this was adjusted from 6.5 to 8.5 with 10 M NaOH. Q Sepharose FF resin (purchased from Pharmacia LKB Biotechnology) was equilibrated with ice cold 20 mM Tris.HCl, pH8.5 and made to a 50% slurry with this buffer. 20 mls of the slurry was added to the diluted filtrate in a 2 l glass beaker. The suspension was stirred gently for 20 minutes on ice. The resin was then packed into a column and washed with 4 column volumes of ice cold 20 mM Tris.HCl, pH8.5. Packing resin into a small diameter column means that proteins can be eluted in small volumes. Bound proteins were eluted from the matrix using 20 mM Tris.HCl, pH8.5, 1 M NaCl. A pale brown band was seen to traverse down the column and was collected. A chromatogram of the fractions together with their relative protease activities, which were also monitored, is shown in Figure 5.8.

It appeared that this type of batch binding and step elution would form an easy and practicable first step in obtaining comparatively high levels of Ara12 protease from a dilute, but plentiful source. This was indeed later incorporated as a first step in the overall purification strategy used to purify Ara12 protease to homogeneity.

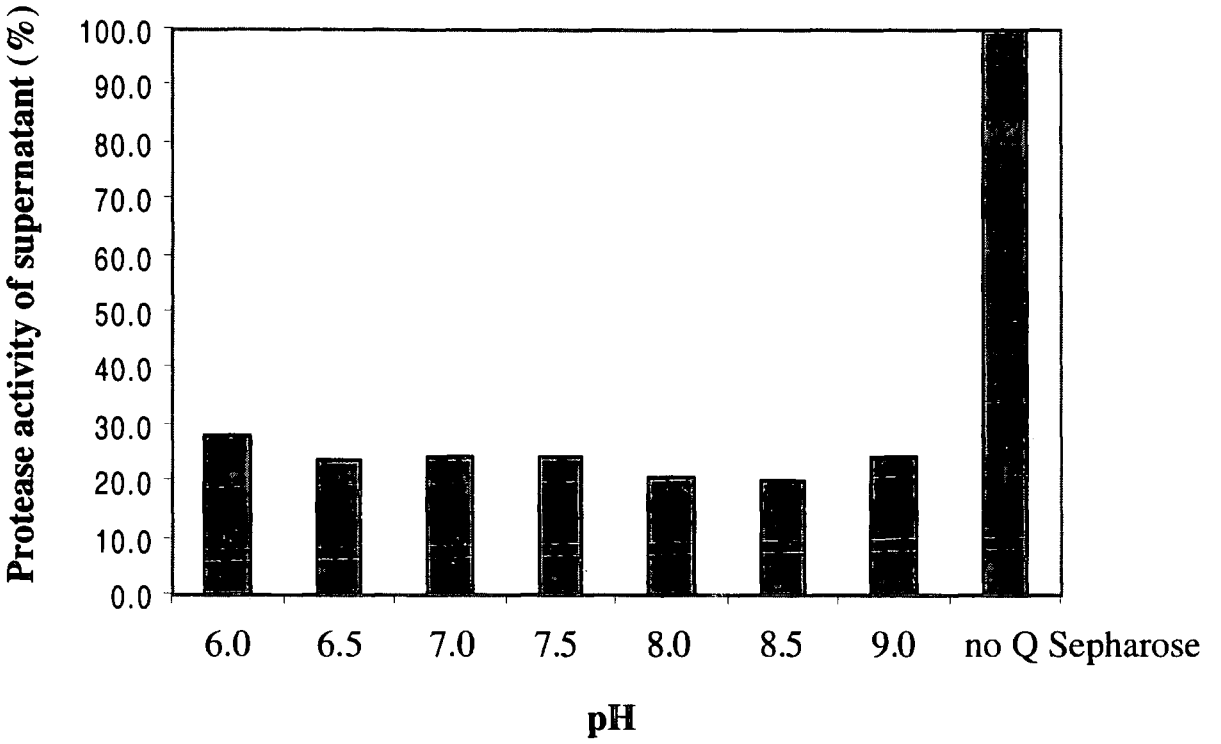


Figure 5.7 The effect of pH on the binding of Ara12 protease to Q Sepharose FF resin. This was monitored by a reduction in protease activity of *Arabidopsis* culture ultrafiltrate (60x ACF) with addition of Q Sepharose FF resin, which had been equilibrated with 10mM Tris.HCl adjusted to pH6.0, 6.5, 7.0, 7.5, 8.0, 8.5 or 9.0. The mixture was gently shaken for 40 minutes. No resin was added to a control tube which had a final pH of 7.5. Samples were then centifuged to pellet the resin. Protease activity of the supernatants was measured with FTC-casein. Least protease activity was detected using resin which had been equilibrated with 10mM Tris.HCl, pH8.5

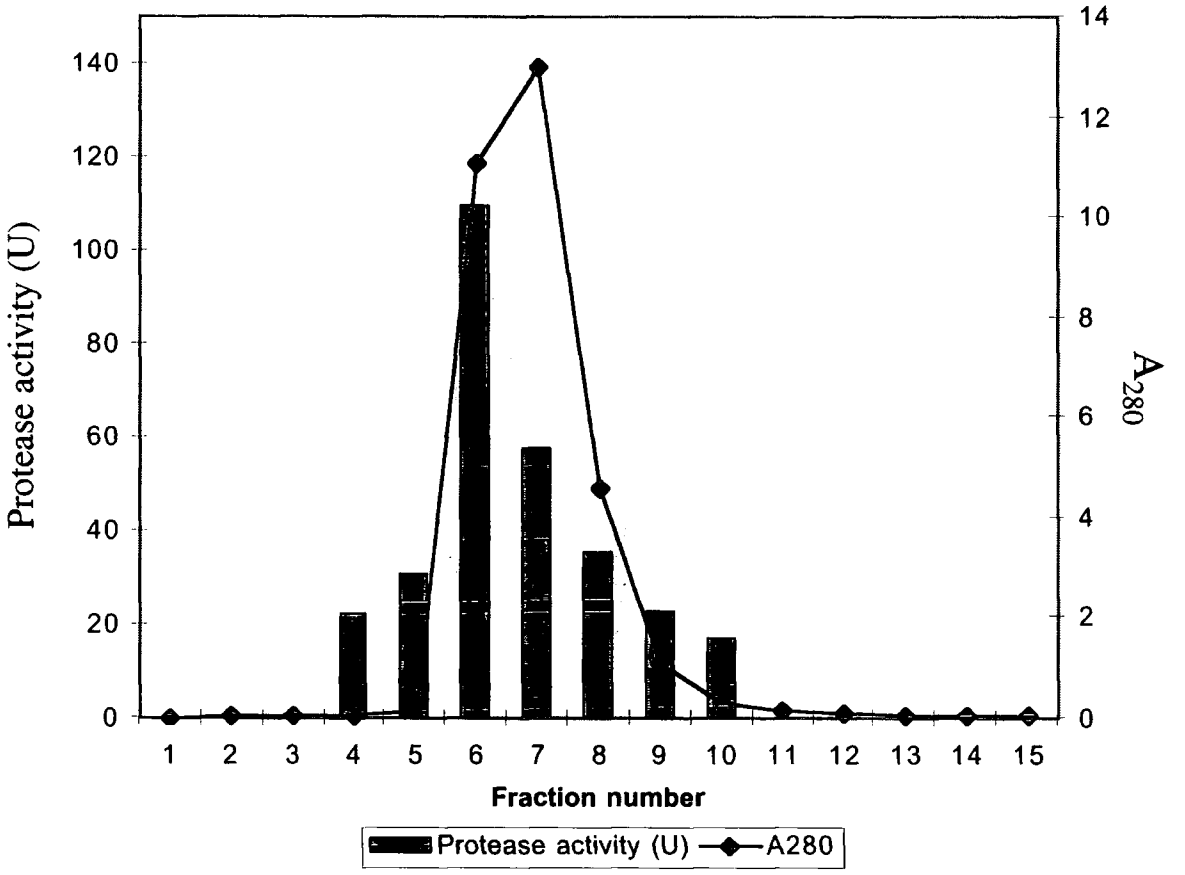


Figure 5.8 Elution profile of *Arabidopsis* culture filtrate proteins batch bound to Q Sepharose FF ion exchange resin. Proteins were step eluted with 1M NaCl, 20mM Tris.HCl, pH8.5 and 1ml fractions were collected. Protease activity was measured using FTC-casein and is represented by the bars. One unit (1U) of FTC-casein degrading activity is defined as the amount of enzyme required to produce 1.0 unit of Ex485 Em538 fluorescence increase under the standard conditions of the assay. A measure of protein concentration, the absorbance at 280nm (A_{280}), is shown as a line with dark diamonds.

An additional batch binding experiment of ACF proteins to Q Sepharose FF was performed, however different step elution conditions were tested to ascertain under which set of conditions the protease activity associated with the filtrate would be eluted. 700 mls of freshly filtered ACF, obtained six days after subculture, was diluted to a total volume of 5 l with cold Milli Q water. As a result the conductivity dropped from 17.2 mS/cm to 0.45 mS/cm. The pH of the solution was adjusted to pH8.5 and 55 mls of a 50% slurry of Q Sepharose FF equilibrated with 20 mM Tris.HCl, pH8.5 was added. The mixture was stirred for 1 h at 4°C, before packing the resin into a 25/20 column. The column was washed with 4 column volumes of the equilibration buffer. Bound proteins were step eluted from the matrix with 0.1 M NaCl, 0.2 M NaCl and 1 M NaCl. A chromatogram showing the elution profile and protease activities of 1.6 ml fractions obtained is shown in Figure 5.9. This showed that whilst the vast bulk of the protein displaying protease activity was eluted in 0.1 M NaCl only a fraction of the total protein was eluted at this ionic strength. This observation was later applied in the first step of the purification of Ara12 protease.

5.5 Selection of the starting material: ACF or silique tissue?

From Western blot hybridisation analysis it was established that immature silique tissue is comparatively rich in Ara12 protease (see Figure 4.6). An investigation was undertaken to evaluate the suitability of silique tissue as a starting material from which to purify functional Ara12 protease. A silique extract was generated and used in ion exchange

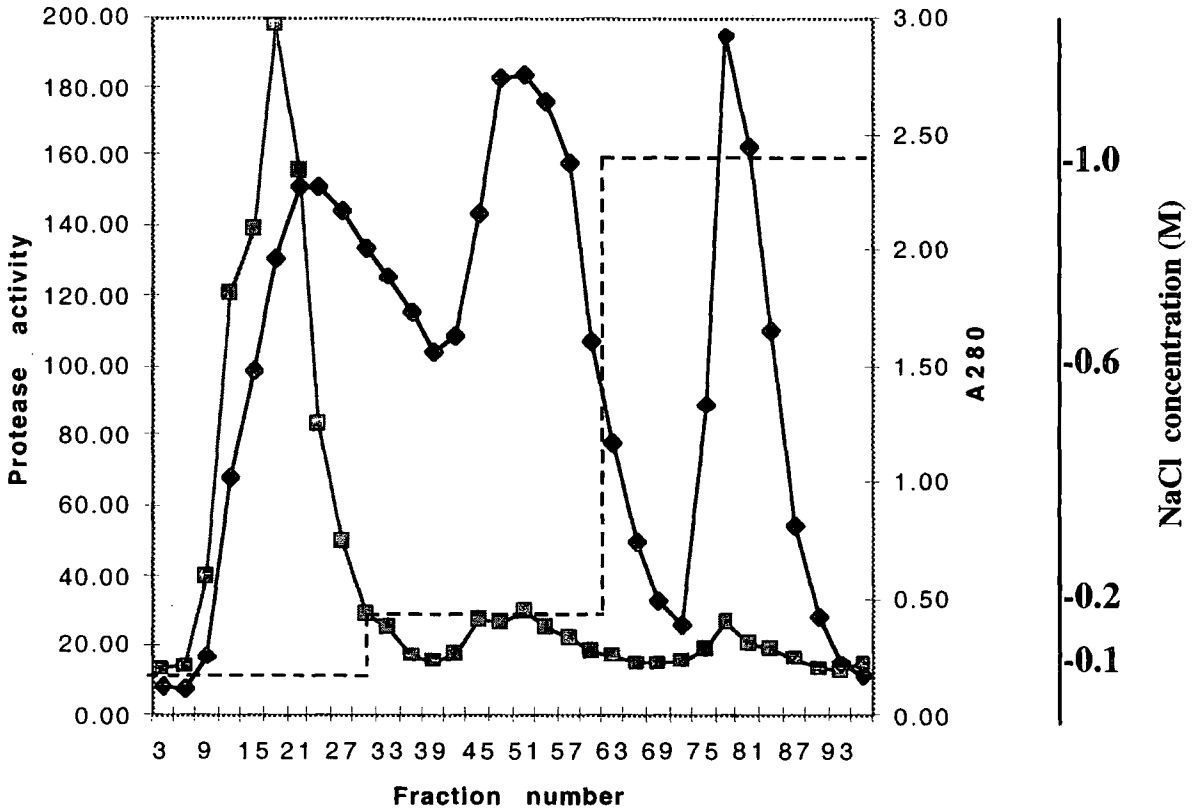


Figure 5.9 Differential step elution profile of ACF proteins batch bound to Q Sepharose FF resin. Growing *Arabidopsis* suspension cultured cells were filtered six days after subculture. The filtrate was diluted with deionised water and the pH was adjusted to 8.5. Anion exchange resin was mixed with this solution for 1 hour. Proteins bound to the resin were step eluted with 0.1, 0.2 and 1.0M NaCl and 1.6ml fractions were collected. The NaCl concentration is represented by the dotted line. Fractions were assayed for proteolytic activity with FTC-casein. Protease activity of the fractions is shown in units by the line with grey squares. The absorbance at 280nm is shown by the line with black diamonds.

chromatography to determine whether Ara12 formed a major component of the protease activity of the extract and to give some idea of how complex the protease activity was.

One gram of immature silique tissue from *Arabidopsis* plants was ground to a fine powder in liquid nitrogen and suspended in 3 ml of ice cold 20 mM Tris.HCl, pH8.5. The suspension was thoroughly mixed and filtered first through four layers of Miracloth and then through a 0.2 μm filter unit. This solution was dialyzed overnight against 20 mM Tris.HCl, pH8.5 at 4°C. One millilitre (0.7 mg) of this solution of silique proteins was loaded onto a 1.6/5 Mono-Q ion exchange column which had been equilibrated with 20 mM Tris.HCl, pH8.5. After loading the column was washed with 4 column volumes of the same buffer. Elution from the column took place over 15 column volumes on a linear gradient of 0-0.4 M NaCl. Twenty four 100 μl fractions were collected and assayed with FTC-casein. A chromatogram with the protease assay results are shown in Figure 5.10. Fractions were not examined by SDS-PAGE or Western blotting, which would have helped to determine which fractions contained Ara12 protease. Unlike the ion exchange chromatography results obtained using ACF as a source of Ara12 protease, no one clear and dominant peak of protease activity was observed amongst the fractions. Compared to ACF, Ara12 protease may be expected to be found as a lower proportion of the total protein in the silique extract. Large quantities of ACF could be generated with ease in less than a week, however it would take a number of weeks to grow *Arabidopsis* plants to the stage at which immature siliques could be harvested (Pang and Meyerowitz, 1987). Furthermore, the silique protein extract would be expected to consist of a much more complex mixture of proteins, making a purification from silique tissue more involved.

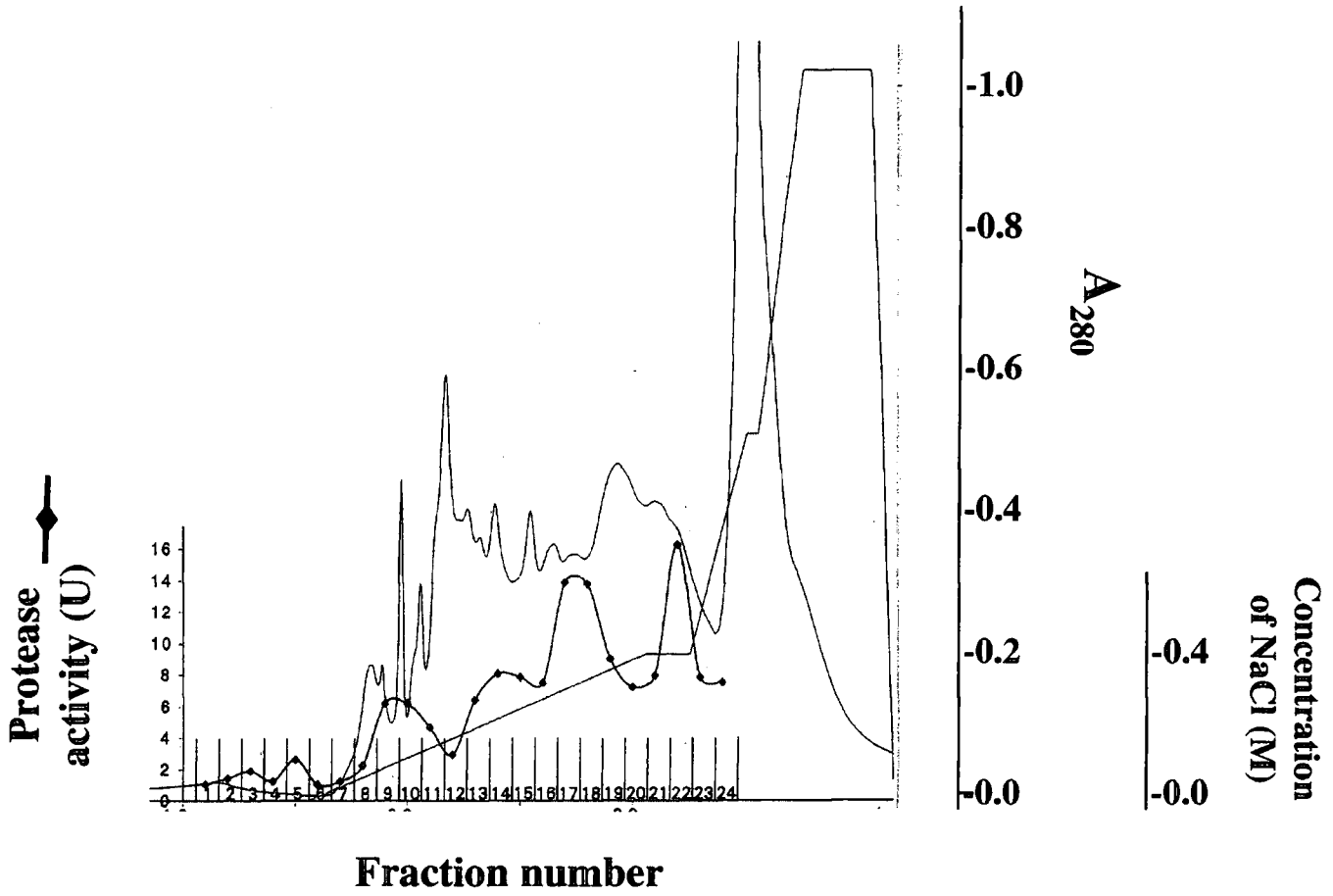


Figure 5.10 Chromatogram of crude *Arabidopsis* silique protein extract fractionated by HPLC on a Mono-Q column. Immature silique proteins were ground to a fine powder and extracted with ice cold 20mM Tris.HCl, pH8.5. 0.7mg of the extract was fractionated on a 1.6/5 Mono-Q column. Proteins were eluted on a linear gradient of 0-0.4M NaCl as shown above. The protease activity of the 100µl fractions was monitored using FTC-casein and is shown by the line with black diamonds. The other line on the chromatogram shows the absorbance at 280nm (A_{280}).

For these reasons it was decided to use the ACF, and not the immature silique tissue, as a source from which to attempt the purification of Ara12. Not only was the filtrate rich in active Ara12 protease, but it also had a relatively low total protein content.

5.6 A search for other possible purification routes

A number of other purification steps were assessed for their suitability for inclusion in a purification strategy. These were mainly chromatographic steps (using hydroxylapatite resin, the affinity chromatography resins benzamidine and arginine Sepharose, the cation exchange resin Mono-S and the hydrophobic interaction chromatography resin phenyl Superose). An immunoprecipitation approach was also examined.

5.6.1 Hydroxylapatite chromatography

A batch binding experiment was carried out using hydroxylapatite and *Arabidopsis* culture filtrate (ACF), which had been dialyzed against 20 mM Tris.HCl, pH6.8, however the results demonstrated that under the conditions used on average only just over 40% of the apparent protease activity could be removed from the supernatant with the resin (data not shown). Efficient protein adsorption to hydroxylapatite occurs only at low ionic strengths and, although care was taken to keep the conductivity of the sample to a minimum, an insufficient amount of proteolytically active protein could be extracted from the filtrate to merit inclusion as a step in the purification. An equivalent batch binding

experiment using Q Sepharose FF resin resulted in the removal of about 80% of the apparent protease activity from culture filtrate (see Figure 5.7).

5.6.2 Immunoprecipitation

Two different immunoprecipitation experiments using 20x ACF and polyclonal anti-Ara12 antibodies were carried out. This method of purifying proteins relies on possessing an antibody with high affinity for the protein of interest and using it to draw that protein out of aqueous solution. In both experiments anti-Ara12 serum (generated during the course of this work) was added to concentrated and dialyzed *Arabidopsis* culture filtrate. In the first procedure, biotinylated anti-rabbit IgG was then added, followed by magnetic beads (Dynabeads) which were bound to streptavidin. After incubation, in theory, the Ara12 protein should bind to anti-Ara12 serum, which should bind to the secondary antiserum, which in turn should bind to the Dynabeads via the biotin-streptavidin interaction. In this way it would be feasible to literally pull Ara12 protein out of solution with a magnet. A second experiment involved adding protein A-Sepharose CL-4B resin to 20x culture filtrate, which had been allowed to incubate with anti-Ara12 serum. The protein A moiety of the resin binds to IgG molecules and so would bind to anti-Ara12 serum, which in turn had bound Ara12. So Ara12 could be pulled out of solution and be eluted from the resin after removal of the supernatant. Initial trials of these techniques were, however, not successful enough to incorporate into an overall purification strategy and these experiments will not be discussed in more detail.

5.6.3 Affinity chromatography using benzamidine and arginine Sepharose and casein agarose

Two Sepharose resins which have been used in the past to affinity purify serine proteases were examined for their utility in the purification of Ara12 protease. Arginine Sepharose 4B (purchased from Pharmacia) has been used in the past to purify an extracellular matrix degrading trypsin-like serine protease (Simon *et al.*,1987). A batch binding experiment and a column separation were performed with concentrated culture filtrate using this resin. ACF was Amicon concentrated 60- fold and 100 µl of this solution was placed into Eppendorf tubes with 200 µl of a 50% slurry of arginine-Sepharose 4B in 10 mM Tris.HCl at varying pHs (6.5, 7.0, 7.5, 8.0, 8.5 and 9.0). The tubes were incubated at room temperature for one hour, before centrifuging for 2 minutes at 13,000 x g. The supernatant was then assayed for protease activity using the FTC-casein assay. The highest levels of binding of Ara12 to the resin were observed at pH7.5 and 8.0. However approximately 65% of the protease activity still appeared to remain in the supernatant. This may be because Ara12 protease does not cleave a substrate containing arginine, eg. benzoyl arginine p-nitro aniline (BApNA) and therefore may not be expected to bind tightly to arginine residues conjugated to Sepharose resin. It was found that Ara12 protease did not cause detectable levels of BApNA cleavage at 20°C over a 2 h period (data not shown).

The conclusion from this preliminary experiment was that Ara12 could be extracted from ACF using arginine Sepharose, and only a low proportion of the total available bound to the resin. It was decided to adhere to other routes of purification.

A second resin which has been used to purify a calcium-dependent serine protease from yeast α-cells (Mizuno *et al.*, 1987), benzamidine Sepharose 4B, was purchased from

Pharmacia and used in batch binding experiments with ACF concentrated 60-fold. p-amidobenzamidine Sepharose 4B was equilibrated and made up to a 50% slurry with solutions of 10 mM Tris.HCl of varying pH (6.0, 6.5, 7.0, 7.5, 8.0, 8.5 and 9.0). 200 µl of this slurry was added to 150 µl of 60x ACF which had been dialyzed against 10 mM Tris.HCl of the corresponding pH. Tubes were left shaking gently at room temperature for one hour, before centrifuging the samples at 13,000 x g for 2 minutes and removing the supernatant. Pellets were resuspended in 80 µl of 10 mM Tris.HCl (variable pH in accordance with that of the slurry) and 20 µl 5 M NaCl. After a 10 minute incubation the tube contents were spun down as before and the eluent was removed. Proteins bound to the resin should elute under the elution conditions described (Mizuno *et al.*, 1987). Eluents from the tubes were assayed using FTC-casein, along with the load and resuspended pellet material from the experiment. Approximately 60% of the protease activity remained in the supernatant and did not bind to the matrix under the conditions described. Due to the relatively small proportion of protease activity which could be removed from the 60x ACF protein sample, no fractionation of the 60x ACF protein sample was attempted using benzamidine Sepharose and this affinity resin was not considered any further for the purification of Ara12.

A number of proteases have been affinity purified using a resin with a protein conjugated to it. The principle of this type of affinity purification is that most of the proteins in the crude extracts will pass through a column containing the gel matrix, however proteases tend to be adsorbed onto the column due to their affinity for other proteins. Adsorbed proteins can be eluted by altering the pH or conductivity of the eluent. Examples of this include the purification of two calcium-dependent proteases using casein Sepharose

(DeMartino and Croall, 1983). Haemoglobin agarose was used to isolate cathepsin D from bovine spleen and thymus (Smith and Turk, 1974) and haemoglobin Sepharose was used to purify proteases from malted wheat flour (Chua and Bushuk, 1969). From previous results shown in this chapter it was demonstrated that protease activity can be monitored in ACF using a fluorogenically labelled casein substrate. Therefore α-casein covalently linked to agarose (purchased from Sigma) was used in a small scale batch binding experiment to determine whether the protease activity found in 60x ACF could be removed by this gel matrix.

α-casein agarose was equilibrated and made up to a 50% slurry with solutions of 10 mM Tris.HCl of varying pH (6.5, 7.5 and 8.5). 200 µl of this slurry was added to 400 µl of 60x ACF which had been dialyzed against 10 mM Tris.HCl of the corresponding pH. Tubes were left shaking gently at room temperature for one hour, before centrifuging the samples at 13,000 x g for 2 minutes. The relative protease activities of the supernatants were determined. From these results the matrix appeared to bind 38% of the protease activity at pH6.5, but at higher pH values less binding was observed.

It appeared that the affinity chromatography steps examined here could not readily be used to efficiently purify Ara12 protease, because only a small proportion of the protease activity was extractable.

5.6.4 Cation exchange chromatography with Mono-S

The following chromatographic procedure was undertaken to see the viability of using cation exchange chromatography as a step in the purification. The isoelectric point (pI) of

mature Ara12 protease has been calculated to be 6.4, and thus this protein has a net positive charge. Therefore it could be advantageous to use a negatively charged (cat)ion exchange resin, such as Mono-S. A portion of the pooled proteolytically active fractions from the batch binding and elution from Q Sepharose FF (shown in Figure 5.8), was dialyzed overnight against 20 mM sodium acetate, pH5.0 at 4°C. One millilitre of the resulting solution was loaded onto a 1.6/5 Mono-S (sulphonate) column, which had been equilibrated with the same buffer used in the dialysis step. The column was then washed with 4 column volumes of this buffer. Elution took place with a linear gradient of 0-0.4 M NaCl in 20 mM sodium acetate, pH5.0 over 15 column volumes. One hundred microlitre fractions were collected and assayed for protease activity. The resultant chromatogram and the corresponding relative protease activities demonstrated that a Mono-S fractionation could effectively resolve the proteolytically active protein in the mixture (data not shown). The protease was eluted from the Mono-S column in approximately 0.2 M NaCl. This step was not used in the final purification.

5.6.5 Hydrophobic interaction chromatography with phenyl Superose – a possible second step in the purification of Ara12 protease

Hydrophobic interaction chromatography (HIC) relies on the hydrophobic interaction which can occur in aqueous solutions between proteins and matrices with hydrophobic ligands, i.e. alkyl or aryl side chains (Hjerten *et al.*, 1974). The interaction is elevated at high ionic strengths and protein desorption occurs from HIC gel matrices as the ionic

strength is reduced. This technique can afford a high degree of purification with a good recovery of the loaded protein.

An equal volume of 3.6 M $(NH_4)_2SO_4$, 20 mM Tris.HCl, pH8.5 was added to a portion of the most proteolytically active fraction from the batch binding using Q Sepharose FF (fraction 17 from the elution shown in Figure 5.9). A phenyl Superose column (1.6/5) was equilibrated with 1.8 M $(NH_4)_2SO_4$, 20 mM Tris.HCl, pH8.5 at 4°C. One millilitre of the sample was loaded onto the column at 200 µl/min. After loading, the column was washed with the equilibration buffer used. Protein bound to the column was eluted over 15 column volumes with a linear gradient of 1.8-0 M $(NH_4)_2SO_4$ in 20 mM Tris.HCl, pH 8.5 and 100 µl fractions were collected. A chromatogram is shown together with the protease activities determined using FTC-casein (see Figure 5.11). The active fractions were analysed by SDS-PAGE as shown in Figures 5.12 and 5.13. This hydrophobic interaction chromatography step was subsequently used in the final purification of Ara12 protease.

In order to determine whether the protease activity found in the active fractions was caused wholly or in part by one protein of the correct predicted molecular weight, a gel slice assay was performed using this partially purified sample (this sample had been batch bound and eluted from Q Sepharose FF and then fractionated on phenyl Superose).

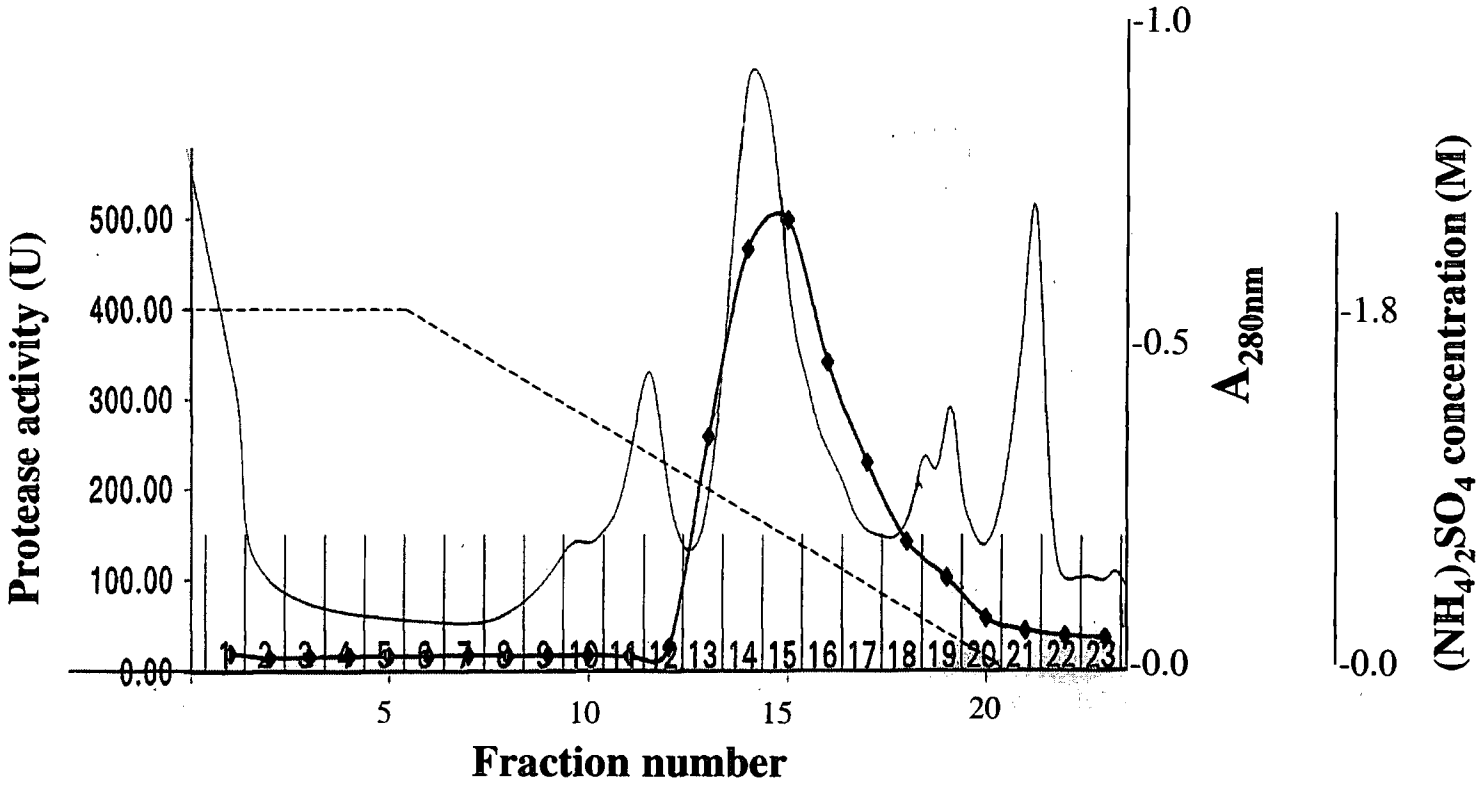


Figure 5.11 Chromatogram of *Arabidopsis* culture filtrate proteins batch bound to Q Sepharose and then fractionated by HPLC on a phenyl Superose column. An equal volume of 3.6M $(NH_4)_2SO_4$ was added to the most proteolytically active fraction (fraction 17) from the Q Sepharose FF batch elution shown in Figure 5.9. This sample was loaded onto a 1.6/5 phenyl Superose column. Proteins were eluted on a linear gradient of 1.8-0M $(NH_4)_2SO_4$ shown by the dotted line. The protease activity of the 100µl fractions was determined using FTC-casein and is shown by the line with the black diamonds. The other line on the chromatogram shows the absorbance at 280nm (A_{280nm}).

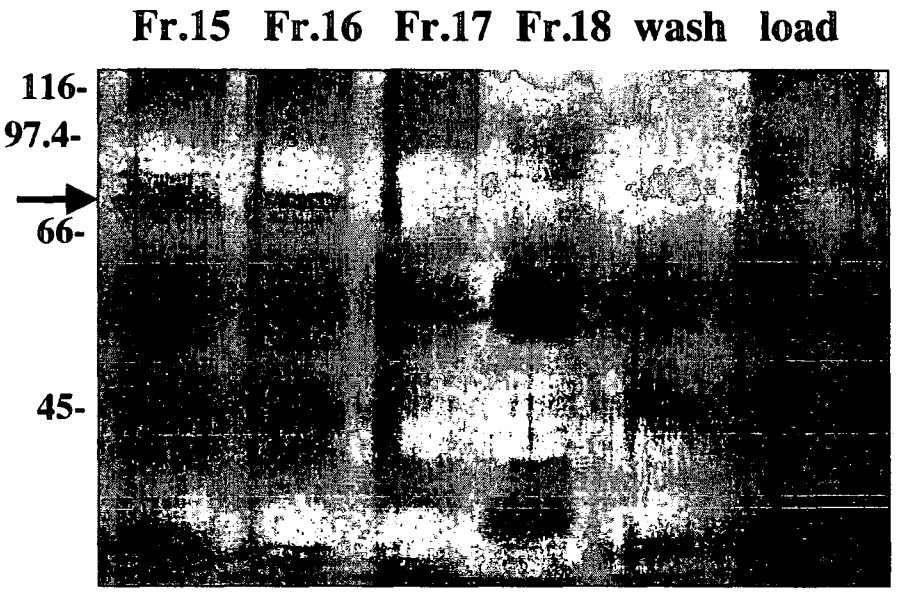


Figure 5.12 SDS-PAGE analysis of *Arabidopsis* cell culture ultrafiltrate fractionated using phenyl Superose. The load (60x ACF), wash and fractions with different levels of protease activity from the phenyl Superose fractionation were resolved on an 8% polyacrylamide gel. 10µl of the fractions indicated, 10µl of the wash and 10µl of the load were resolved on the gel. Proteins were visualized by silver staining. The gel appears slightly hazy, because the samples resolved were not desalted before running. Molecular masses are given in kDa. Fraction 15 showed the highest protease activity levels in this fractionation. Mature Ara12 protease is indicated with an arrow.

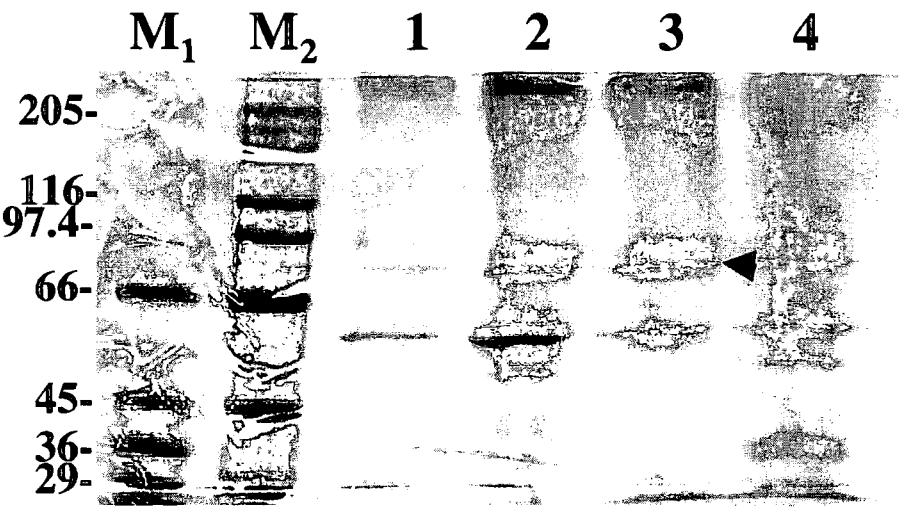


Figure 5.13 SDS PAGE analysis of partially purified Ara12 protease. *Arabidopsis* culture filtrate (ACF) proteins were batch bound to Q Sepharose FF resin at 4°C and step eluted in 0.1M NaCl (as shown in Figure 5.9). The eluate fraction with the highest protease activity (fraction 17) was fractionated on a 1.6/5 phenyl Superose column (see Figure 5.11). Protein samples from each stage were resolved on an 8% polyacrylamide gel. Lane 1, 20µl ACF (undiluted and unconcentrated); lane 2, 20µl Q batch fraction 17; lanes 3-4, 20µl fractions 15 and 16, respectively, from the phenyl Superose HPLC separation. Fraction 15 showed the highest specific activity of the samples examined here. Proteins were visualized by Coomassie staining. The protein molecular weight markers SDS7 (M_1) and Hi-6 (M_2) used were purchased from Sigma and are identified in kDa. Mature Ara12 protease can be seen in lanes 2-4 and is indicated with an arrowhead.

5.6.6 Gel slice assay of partially purified Ara12 protease

Partially purified Ara12 protease obtained from the phenyl Superose fractionation (fraction 15) described in the previous section was run on a native 8% polyacrylamide resolving mini-gel (lacking SDS), which had no stacker gel. The sample was loaded in native loading buffer (without DTT and SDS) and was not boiled. Proteins were resolved at 15 mA at 4°C in 1x PAGE running buffer lacking SDS. Three main smeary bands were visible on the native gel by Coomassie staining as shown in Figure 5.14 A). An identical lane of the gel was not stained and cut into 2 mm strips perpendicularly to the direction of protein migration. Each strip was placed into a separate Eppendorf tube and ground up in 100 µl of 10 mM Tris.HCl, pH7.5, 6.25 mM $CaCl_2$ (homogenisation buffer) with a homogeniser. The tubes were centrifuged at 13,000 x g for 5 minutes, before assaying the supernatant for protease activity with FTC-casein. The protease assay results are shown in Figure 5.14 A) beside the native gel and the activity appeared to be found mainly in gel slices 4 and 5. The extracts of these two gel slices were subsequently analysed by conventional SDS-PAGE and Coomassie staining to determine which protein(s) they contained (see Figure 5.14 B)). Only one protein was visible in the two most active gel slice extracts. The apparent molecular weight of this protein was consistent with the protein being Ara12. This suggested that the major protease activity witnessed was, as far as could be detected, wholly caused by one protein, likely to be Ara12 subtilisin-like protease.

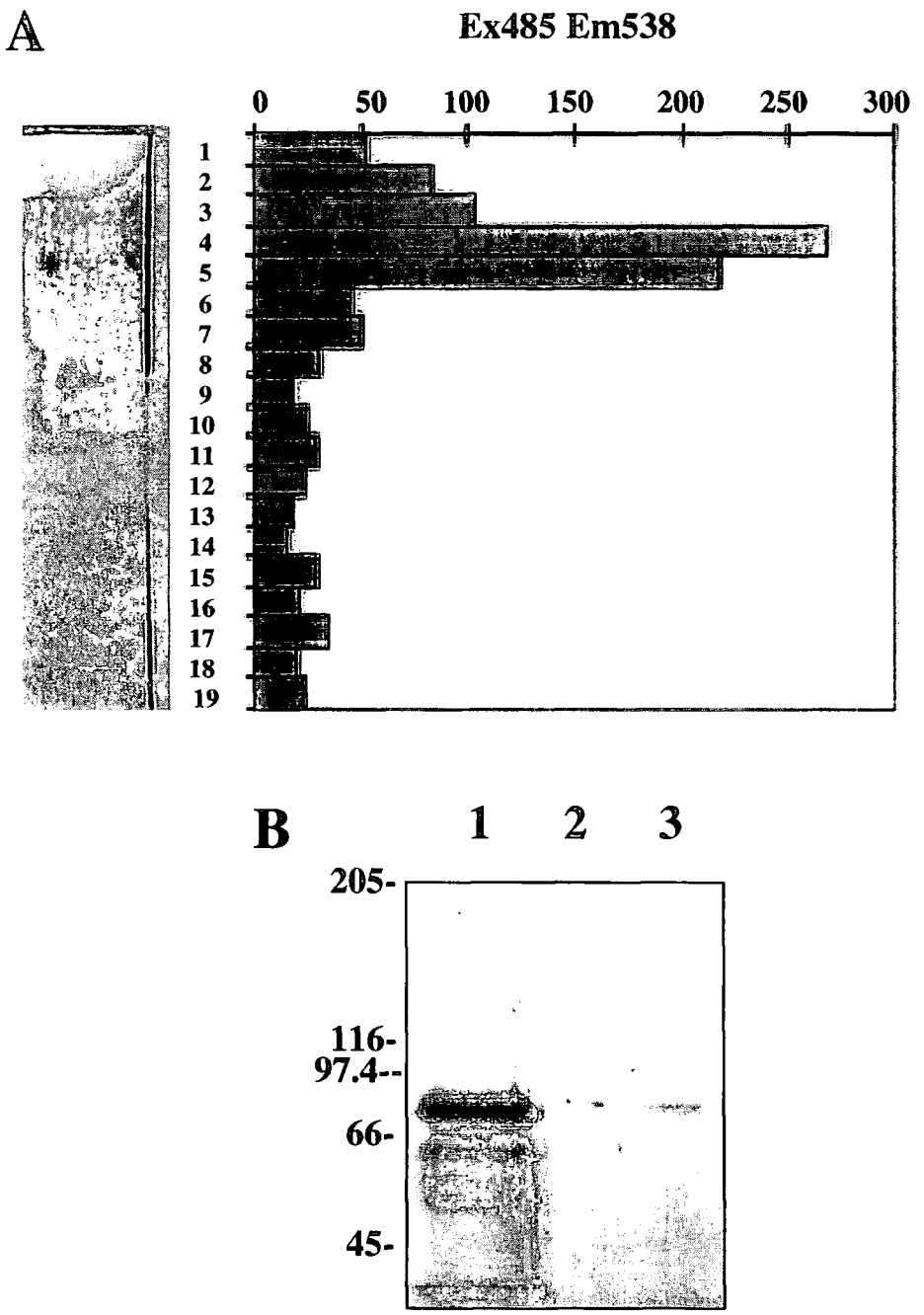


Figure 5.14 **A)** Gel slice protease assay of the partially purified Ara12 sample generated by differential Q Sepharose batch elution followed by fractionation on phenyl Superose. Proteins were separated on an 8% native polyacrylamide gel and the gel was cut into 2mm strips. Protein was extracted from the gel slices with 10mM Tris.HCl, pH7.5; 6.25mM $CaCl_2$ and protease activities of the extracts were determined with FTC-casein. **B)** Extracts displaying relatively high protease activities were analysed by SDS-PAGE on an 8% polyacrylamide gel. Lane 1, 10µl of partially purified Ara12 sample as applied to gel; lane 2, 10µl extract fom gel slice number 4; lane 3, 10µl extract from gel slice number 5.

5.7 Purification of Ara12 subtilisin-like protease

This section describes the purification of Ara12 protease from *Arabidopsis* suspension culture filtrate in four steps: 1) batch binding to and selective step elution from Q Sepharose FF, 2) binding to and step elution from phenyl Sepharose FF, 3) fractionation on phenyl Superose and 4) fractionation on Mono-Q. The purification procedure is outlined in Figure 5.15.

5.7.1 Batch binding of *Arabidopsis* suspension culture filtrate proteins to Q Sepharose FF – purification step 1

Six litres of 6 day old *Arabidopsis* suspension culture filtrate was diluted with ice cold Milli Q water to fifteen litres. The conductivity was reduced in this way from 3.8 mS/cm to 1.5 mS/cm. Before dilution the pH had been adjusted to 8.5 using 10 M NaOH. 650 ml of 50% Q Sepharose FF slurry (obtained from Pharmacia LKB Biotechnology), which had been equilibrated with 20 mM Tris.HCl, pH8.5, was added to the filtrate and this was left at 4°C stirring gently for two hours. This and all chromatography steps in the purification were carried out at 4°C. The slurry was packed into a 25/120 column and was washed at 2 ml/min with four column volumes of 20 mM Tris.HCl, pH8.5. The column was step eluted with four column volumes of 20 mM Tris.HCl, 0.1 M NaCl, pH8.5. It has previously been shown in this work that at this ionic strength Ara12 protease which was bound to Q Sepharose FF, was eluted, however most of the *Arabidopsis* culture filtrate proteins remain bound to the matrix. 6ml fractions were collected using a Gilson fraction collector. A

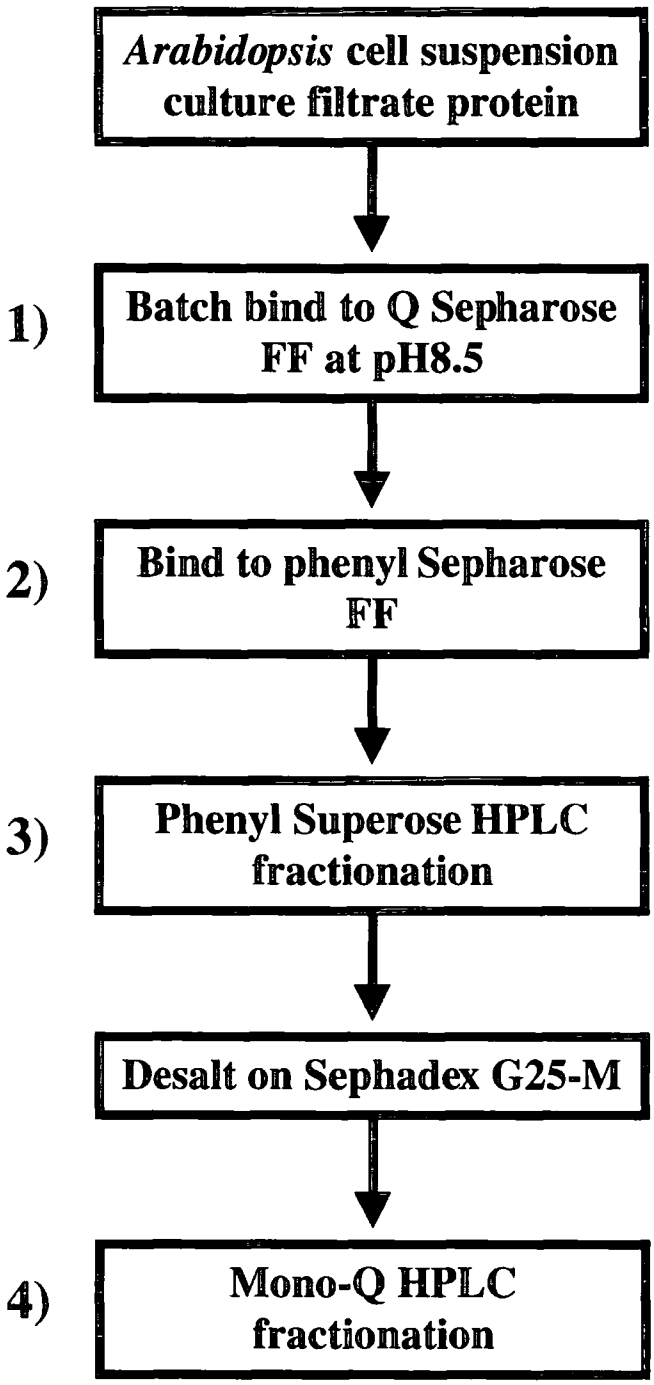


Figure 5.15 Schematic diagram outlining the strategy adopted to purify Ara12 subtilisin-like protease from a source rich in this enzyme: *Arabidopsis* cell suspension culture filtrate. The purification steps 1)-4) are identified here.

brown colour was observed in some of the fractions and these corresponded to fractions which were subsequently shown to have protease activity.

Relative protease activities were determined for the fractions using the FTC-casein fluorimetric assay and are shown with the chromatogram of the Q Sepharose FF batch elution in Figure 5.16. Protease activity was found in fractions 10-14 and these fractions were pooled (to give a total volume of 30 ml) and used in the next purification step.

5.7.2 Binding to phenyl Sepharose FF (FPLC) – purification step 2

An equal volume of 3.6 M $(NH_4)_2SO_4$, 20 mM Tris.HCl, pH8.5 was added to the pooled active fractions (Fr.10-14) from the Q Sepharose FF step elution. This sample was passed through a 2 ml phenyl Sepharose FF column (packed in a 7/15 column), which had been equilibrated with 1.8 M $(NH_4)_2SO_4$, 20 mM Tris.HCl, pH8.5, at a flow rate of 1.4 ml/min. The protein was step eluted with 20 column volumes (40 ml) of 20 mM Tris.HCl, pH8.5 at a flow rate of 1.4 ml/min. This hydrophobic interaction chromatography step was performed using a BioRad BioLogic workstation and 3 ml fractions were collected using a BioRad Model 2128 fraction collector. Fractions 7-9 were found to have protease activity (see Figure 5.17) and were pooled (giving a total volume of 9 ml).

This stage is essentially purely a concentration step using phenyl Sepharose and the eluate was found to have a high ionic strength. The next purification step using a phenyl Superose HPLC column requires the sample to be at a high ionic strength in order to obtain binding of the sample protein to the resin. Instead of adding 3.6 M $(NH_4)_2SO_4$ to the required ionic strength to proceed with the next step, the sample was desalted only to then

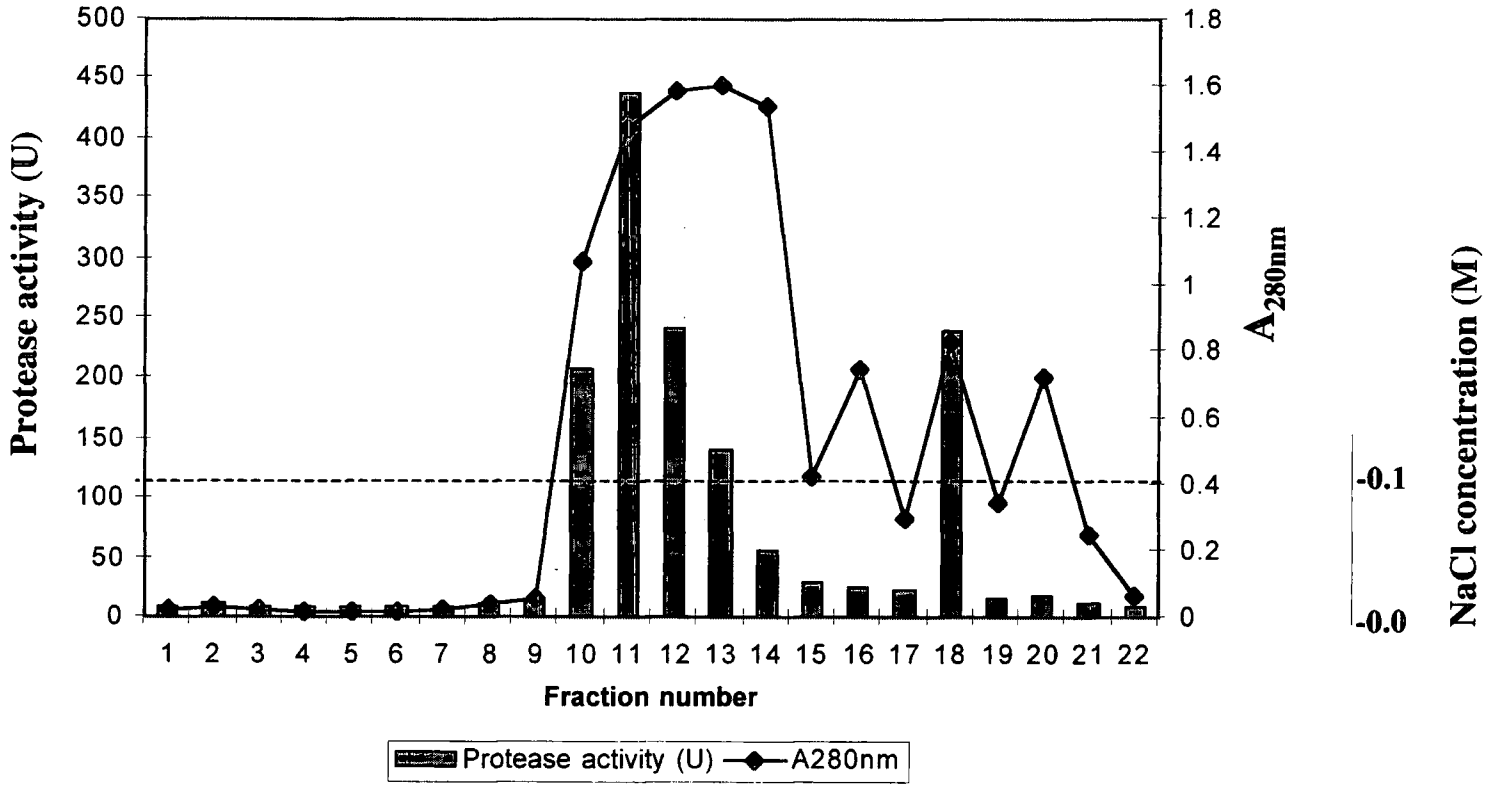


Figure 5.16 Chromatogram of Q Sepharose FF batch elution: Ara12 protease purification step 1. 6l *Arabidopsis* cell suspension culture filtrate was diluted to 15l with Milli Q water and mixed with Q Sepharose FF resin. The slurry was packed into a 25/120 column and washed with four column volumes of 20mM Tris.HCl, pH8.5. Bound protein was step eluted with 0.1M NaCl as shown by the dotted line. Protease activity was determined for the 6ml fractions obtained using FTC-casein and is shown by the bars. The absorbance at 280nm is shown by the line with the black diamonds.

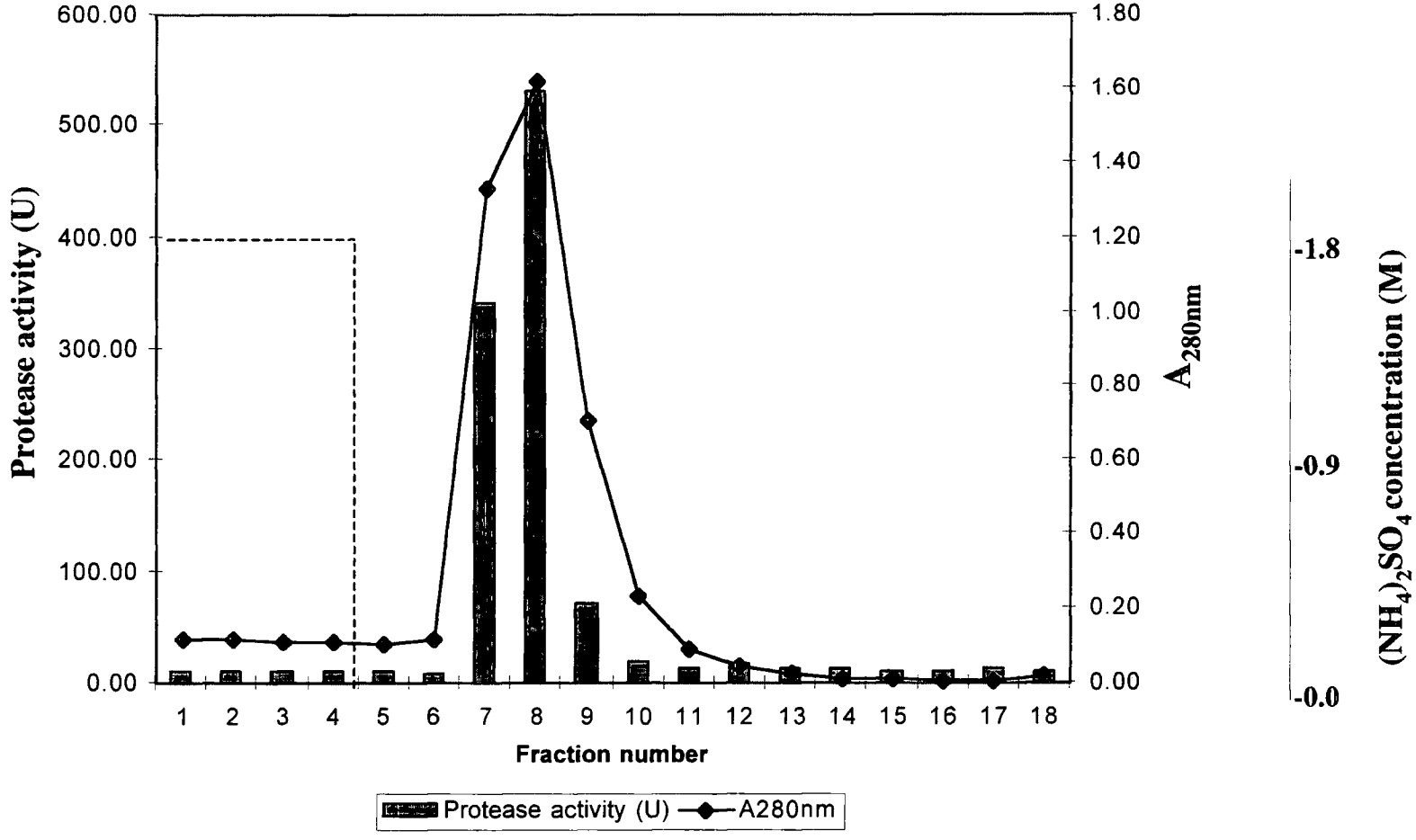


Figure 5.17 Chromatogram of phenyl Sepharose FF elution: Ara12 protease purification step 2. An equal volume of 3.6M $(NH_4)_2SO_4$ was added to the pooled fractions showing proteolytic activity from the Q Sepharose FF batch elution shown in Figure 5.16. This sample was passed down a phenyl Sepharose FF column. Bound protein was step eluted with 20mM Tris.HCl, pH8.5 in the absence of ammonium sulphate as shown by the dotted line. Protease activity was determined for the 3ml fractions obtained using FTC-casein and is shown by the bars. The absorbance at 280nm is shown by the line with the black diamonds.

be brought to 1.8 M $(NH_4)_2SO_4$ by addition of an equal volume of 3.6 M $(NH_4)_2SO_4$. Although on reflection this was an inefficient way to proceed between these two steps, because the sample protein had been unnecessarily diluted, it had little bearing on the final purification.

The volume of the sample was made up to 10 ml with Milli Q water and passed down a PD-10 desalting column (containing Sephadex G-25M resin purchased from Amersham Pharmacia) which had been equilibrated with 10 column volumes (25 ml) 20 mM Tris.HCl, pH8.5. The sample was passed down the PD-10 column under gravity and was eluted using 14 ml of 20 mM Tris.HCl, pH8.5.

5.7.3 Phenyl Superose HPLC – purification step 3

The sample obtained from the previous stage was diluted with 12 ml of 3.6 M $(NH_4)_2SO_4$, taking the sample volume to 26 ml. A phenyl Superose PC 1.6/5 HPLC column was equilibrated with 10 column volumes of 1.8 M $(NH_4)_2SO_4$, 20 mM Tris.HCl, pH8.5 in a SMART system microseparation unit (from Amersham Pharmacia Biotech) at a flow rate of 100 µl/min. The maximum practical loading capacity of this phenyl Superose column is given by the manufacturers as 1.5 mg of protein. The concentration of the load (after addition of ammonium sulphate) was found to be approximately 0.2 mg/ml. Therefore the load contained over 5 mg of protein. To fractionate all the protein in this sample using this phenyl Superose column, the hydrophobic interaction chromatography run was repeated four times using the same load volume, flow rate and salt gradient for elution. As the flow rate and gradient was computer controlled is was possible to obtain almost identical elution

profiles for the four runs. The column was washed with four column volumes of Buffer A and the protein was eluted with a linear salt gradient of 1.8 M-0 M $(NH_4)_2SO_4$ over 20 column volumes. The fractions from each stage of elution for all the runs were then pooled and the protease activities and protein concentrations were determined (see Figure 5.18). The protease activity was found in fractions 10-19 (400 µl pooled fractions). These active fractions were pooled and subsequently desalted on a PD-10 column.

The volume of the pooled active fractions was made up to 5 mls with Milli Q water and passed down a PD-10 desalting column (containing Sephadex G-25M resin purchased from Amersham Pharmacia) which had been equilibrated with 25 ml 20 mM Tris.HCl, pH8.5. The sample was passed down the column under gravity as follows. 2.5 ml of the sample was added to the column and allowed to drain and then 3.5 ml of elution buffer was added. The eluate was collected and the steps, including the initial equilibration step, were repeated until all the initial sample had been passed through the desalting column.

As a result the conductivity of the desalted solution dropped from 30.0 mS/cm to 8.3 mS/cm. The conductivity of this solution was still too high to proceed with the next stage. This was because the solution had been eluted in 9 ml (2x4.5 ml), rather than the 7 ml (2x3.5 ml) stated in the instructions, and therefore some of the salt which would normally have been retained on the column had been eluted with the sample. To desalt the sample further the volume was made up to 10 ml with Milli Q water, passed down a PD-10 column and eluted with 14 ml of 20 mM Tris.HCl, pH8.5. The conductivity of the eluate was still found to be too high to proceed to the next purification step, being 2.6 mS/cm.

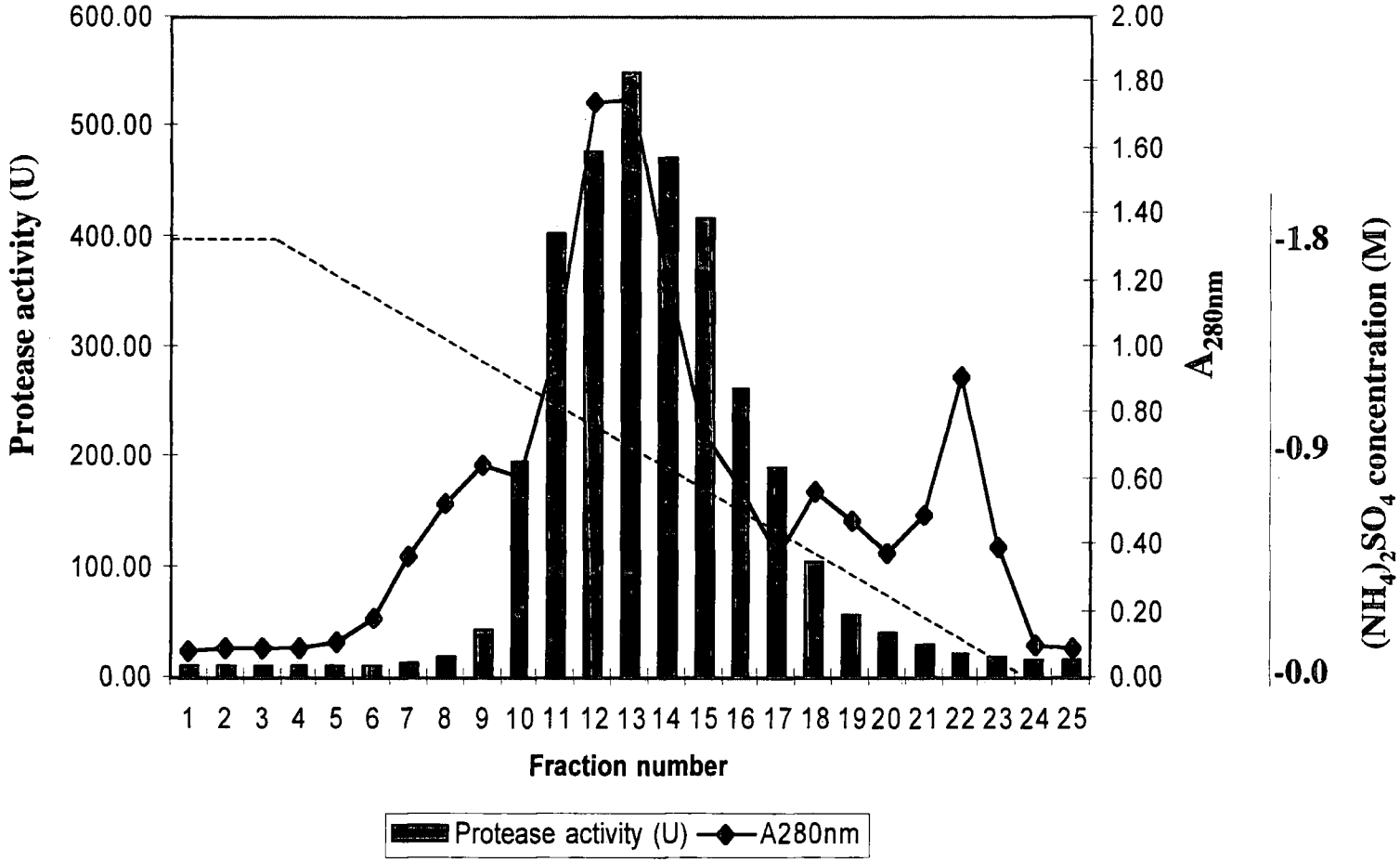


Figure 5.18 Chromatogram of phenyl Superose HPLC fractionation: Ara12 protease purification step 3. Proteolytically active fractions from the phenyl Sepharose elution shown in Figure 5.17 were pooled and applied to a 1.6/5 phenyl Superose HPLC column. The column was washed with 1.8M $(NH_4)_2SO_4$ and bound protein was eluted with a linear gradient of 1.8-0M $(NH_4)_2SO_4$ as shown by the dotted line. Protease activity was determined for the 0.4ml pooled fractions obtained using FTC-casein and is shown by the bars. The absorbance at 280nm is shown by the line with the black diamonds.

Therefore the sample was diluted with Milli Q water to a final volume of 21 ml, giving a final conductivity of 1.6 mS/cm.

5.7.4 Mono-Q anion exchange HPLC – purification step 4

A 1.6/5 Mono-Q ion exchange column was used to fractionate the proteins further by anion exchange chromatography, as a last step in the purification. The column was equilibrated with 10 column volumes of 20 mM Tris.HCl, pH8.5 at a flow rate of 100 μ l/min on a SMART system HPLC microseparation unit. The maximum practical loading capacity for this column is 0.5 mg of protein. The protein concentration of the load at this stage was approximately 0.12 mg/ml, giving a total mass for the load of about 2.5 mg protein. To fractionate all of the sample using this column, seven sequential chromatography runs were performed with 3 ml of the sample per run. After equilibrating and loading the sample manually using a 1 ml loading loop, the column was washed with 4 column volumes of 20 mM Tris.HCl, pH8.5. The run conditions were kept as constant as possible, for example, the same flow rate of 100 μl/min and the same salt gradient for protein elution was used throughout. Protein bound to the column was eluted with a linear gradient of 0-0.1 M NaCl in 20 mM Tris.HCl, pH8.5 over 40 column volumes and 200 μl fractions were collected. All procedures were carried out at 4°C.

As before the fractions from each stage of elution for all the runs were then pooled and the protease activities and protein concentrations were determined (see Figure 5.19). It was found that the protease activity resided mainly in fractions 4-6 (with a volume of 1.4 ml

per fraction) and these active fractions were pooled giving a total volume of 4.2 ml for the purified Ara12 protease fraction.

5.8 Confirmation of the identity of purified Ara12 protease

SDS PAGE analysis of pooled active fractions from each stage of the purification is shown in Figure 5.20. The identity of the purified Ara12 protease was confirmed by immunoblot analysis using anti-Ara12 serum (see Figure 5.21). Also, the N-terminal sequence of the final purified protein sample was determined using an Applied Biosystems 477A protein sequencer (many thanks to John Gilroy for sequencing the protein). The identity of this N-terminal sequence confirmed that the protein purified was indeed the mature Ara12 subtilisin-like protease (the first ten amino acid residues were determined as TTRTPLFLGL). The yield per cycle of the N-terminal sequencing is given in Table 5.1.

The purified protein had a molecular weight just below 80 kDa, however according to the primary sequence of mature Ara12 the molecular weight expected was 67.6 kDa. The Ara12 protease may be glycosylated, accounting for the difference in these molecular weights. Ara12 contains 7 potential glycosylation sites. A digoxigenin (DIG) glycan detection kit (obtained from Boehringer Mannheim) was used to assess whether Ara12 was glycosylated (data not shown). However, the results were inconclusive and could not be interpreted to definitively determine whether or not Ara12 was glycosylated.

Table 5.2 shows the yield and purification of Ara12 protease obtained. From these data the abundance of Ara12 can be determined as 1.9% of the total protein in the culture filtrate.

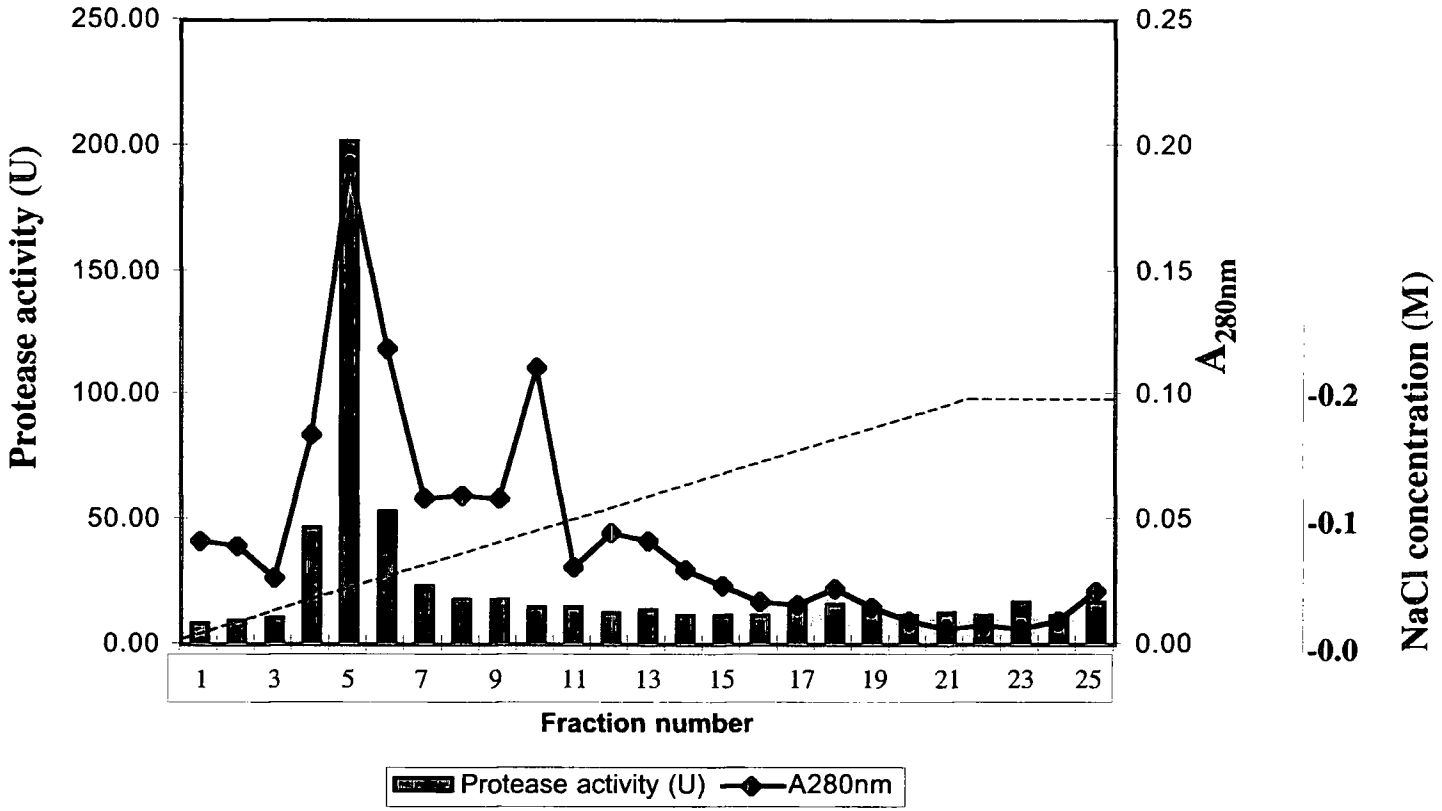


Figure 5.19 Chromatogram of Mono-Q HPLC fractionation: Ara12 protease purification step 4. Proteolytically active fractions from the phenyl Superose fractionation shown in Figure 5.18 were desalted on Sephadex G-25M resin. The sample was applied to a 1.6/5 Mono-Q HPLC column. The column was washed with four column volumes of 20mM Tris.HCl, pH8.5 and bound protein was eluted with a linear gradient of 0-0.2M NaCl as shown by the dotted line. Protease activity was determined for the 1.4ml pooled fractions obtained using FTC-casein and is shown by the bars. The absorbance at 280nm is shown by the line with the black diamonds.

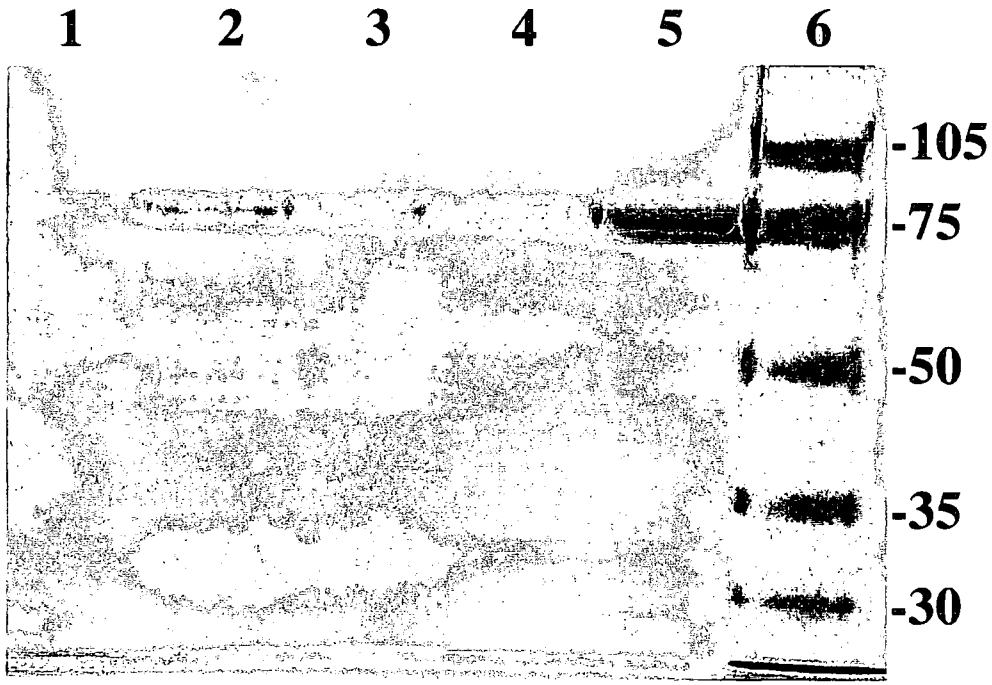


Figure 5.20 SDS-PAGE analysis of the stages of the purification of Ara12 subtilisin-like protease. At each stage of the purification, fractions with protease activity were pooled, desalted if necessary, and 20µl aliquots were resolved on an 8% polyacrylamide gel. Protein was visualized using Coomassie Brilliant Blue R 250 dye. Lane 1, 6d old *Arabidopsis* culture filtrate; lane 2, Q Sepharose FF batch eluate; lane 3, phenyl Sepharose FF batch eluate; lane 4, phenyl Superose fractionation, lane 5, Mono-Q fractionation; lane 6, protein molecular weight markers. Molecular masses are shown in kDa.

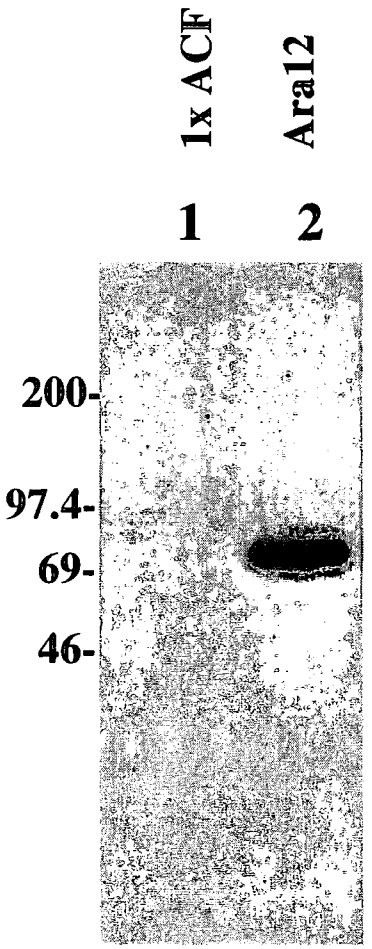


Figure 5.21 Western blot hybridisation analysis of purified Ara12 subtilisin-like protease compared to the *Arabidopsis* culture filtrate (1x ACF) from which it was purified. Lane 1, 1µg of unconcentrated 1x ACF; lane 2, 250ng of purified Ara12 protease. Samples were resolved on an 8% polyacrylamide gel and blotted onto a Hybond-C extra membrane. The membrane was probed with a 1:20,000 dilution of anti-Ara12 serum and a 1:20,000 dilution of goat anti-rabbit IgG-HRP conjugate.

Cycle number/ Amino acid number	Amino acid identity	pmol
1	THR (T)	16.78
2	THR (T)	13.70
3	ARG (R)	18.25
4	THR (T)	10.71
5	PRO (P)	13.18
6	LEU (L)	10.32
7	PHE (F)	9.16
8	LEU (L)	14.22
9	GLY (G)	18.05
10	LEU (L)	14.90

Table 5.1 The yield of the masses of amino acid residues during the determination of the N-terminal sequence of Ara12 protein. The yield is given as picomoles per cycle.

Purification Stage	Protein (mg)	Total Activity (U)	Specific Activity (U/mg protein)	Yield (%)	Purification (-fold)
1) Culture filtrate	390	172,000	441.0	100	—
2) Q Sepharose	7.47	25,348	3393.3	14.7	7.7
3) Phenyl Sepharose	1.90	7797.2	4103.8	4.5	9.3
4) Phenyl Superose	0.74	6339.7	8567.2	3.7	19.4
5) Mono-Q	0.25	5774.1	23,096.4	3.4	52.4

Table 5.2 Isolation of Ara12 subtilisin-like protease from *Arabidopsis* cell suspension culture filtrate. The percentage yield and purification factor of Ara12 subtilisin-like protease is given at the different stages of the purification.

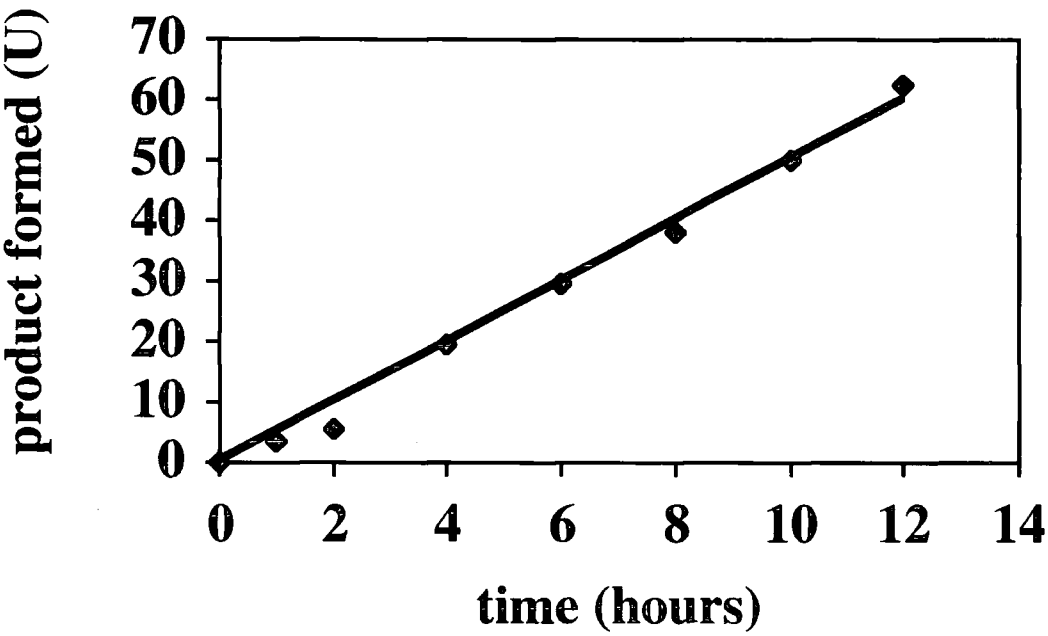


Figure 5.22 Protease activity of Aral2 subtilisin-like protease varying with time. 250ng of purified Aral2 protease was incubated in 0.2% FTC-casein in 20mM Tris.HCl, pH7.5 at 37°C for varying amounts of time, before stopping the 50µl reactions with addition of TCA to a final concentration of 3.5% (v/v). The protease activity was determined from the fluorescence at Ex485 Em538 of TCA soluble peptides. One unit (1U) of FTC-casein degrading activity is defined as the amount of enzyme required to produce 1.0 unit of Ex485 Em538 fluorescence increase under the standard conditions of the assay.

The activity of the purified protease was determined using an assay based on a previously published method demonstrating the degradation of casein labelled with FITC (Twining, 1984) and is shown plotted against time in Figure 5.22. One unit (1 U) of FTC-casein degrading activity is defined as the amount of enzyme required to produce 1.0 unit of Ex485 Em538 fluorescence increase under the standard conditions of the assay.

5.9 Conclusions

The research presented in this chapter has demonstrated the first purification to homogeneity of a subtilisin-like protease from *Arabidopsis thaliana*. Ara12/Slpa protease was purified from a previously characterised source: *Arabidopsis* cell suspension culture filtrate (Robertson *et al.*, 1997). The same cultures used in the study were used in this work. Although another source, immature silique tissue, had been shown to be rich in this enzyme (using a polyclonal anti-Ara12 antiserum), the characterised mixture of extracellular culture filtrate proteins was used as a starting material, from which Ara12 protease was purified to homogeneity.

A purification strategy was devised and implemented, involving anion exchange and hydrophobic interaction chromatography. Q Sepharose and Superose chromatographic steps used in conjunction with phenyl Sepharose and Superose steps were found to be effective methods of purifying Ara12 extracellular protease. Although they were not eventually used in the isolation, a number of other possible ways of isolating the enzyme were explored, such as immunoprecipitation methods, cation exchange and affinity chromatography and hydroxylapatite chromatography. These procedures, with the

exception of Mono-S cation exchange chromatography, did not appear to be very efficient ways of purifying Ara12, under the conditions used.

Initially *Arabidopsis* cell culture filtrate was batch bound to Q Sepharose and eluted in 0.1 M NaCl. The protein was then concentrated on a phenyl Superose column and the eluate was fractionated by HPLC on a phenyl Superose hydrophobic interaction column. Active fractions were, finally, further purified by HPLC on a Mono-Q anion exchange column. The advantage of using small HPLC columns is that a higher degree of resolution can be achieved compared to columns with larger diameters and flow rates.

The Ara12 subtilisin-like protease was purified over 50-fold from *Arabidopsis* cell culture filtrate, a source already rich in this secreted protein. The molecular weight of the protein isolated was larger than that expected for the mature Ara12 protein from its cDNA sequence (67.6 kDa). This could be due to glycosylation of Ara12 and the protein does contain a number of possible glycosylation sites. To determine whether this is the case the protein could be treated with anhydrous hydrogen fluoride, which deglycosylates glycoproteins (Mort and Lamport, 1977). The purified protein showed proteolytic activity against casein substrates (FTC-casein and azocasein). Finally, the identity of the purified protease was confirmed by N-terminal sequencing and immunoblot analysis (using anti-Ara12 polyclonal antibodies).

As is discussed in the next chapter, the purified protease was used to determine the pH and temperature range of activity and its requirement for cofactor ions. Studies were also undertaken to determine the substrate specificity of the enzyme (using extracellular protein extracts from *Arabidopsis* cells, as well as other proteins and synthetic peptides) and to identify inhibitors of its proteolytic activity.

Chapter 6

Biochemical analysis of Ara12/Slpa protease

6.1 Introduction

A number of plant cysteine and aspartic proteases thought to be involved in the proteolytic degradation of storage proteins during seed germination (Cervantes *et al.*, 1994; Mutlu *et al.*, 1999) or regulation of programmed cell death (Solomon *et al.*, 1999) have been characterised. Less, however, is known about plant serine proteases, such as the subtilisin-like proteases. A subtilisin-like protease has not been purified from *Arabidopsis* before. The Ara12 protease, which appears to be found in the extracellular matrix of this well studied plant, and is found at high levels secreted into the culture medium of supension cultured cells, has not been investigated at the biochemical level. The aim of this chapter was to use the Ara12 protease isolated, as described in the previous chapter, in biochemical analysis. There is evidence for a large number of subtilisin-like proteases in *Arabidopsis* (possibly more than fifty) and it seems likely that they control some fundamental processes in plant development and could be involved in processes such as morphogenesis, pathogenesis or housekeeping. Considering the large number of subtilases anticipated just in this plant, basic information regarding the pH and temperature optima and ranges of stability, cofactor requirements and substrate specificities is lacking.

Binding of calcium ions to specific sites can afford stability to proteases by reducing molecule flexibility, which prevents thermal denaturation and autolysis (Siezen *et al.*, 1991). Three dimensional structure studies of subtilisin-like proteases (using X-ray crystallography) have revealed the presence of up to three calcium ion binding pockets in these monomeric enzymes (Smith *et al.*, 1999), and it appears that calcium ion binding has a stabilizing effect on their activity. Therefore it would be of interest to look at possible

cofactors of Ara12 protease, such as calcium and magnesium ions, and, also, how the action of the enzyme would be affected by chelators of these cations, such as the salts of EDTA or EGTA.

Information on the substrate specificity of Ara12 would be useful in determining whether this enzyme acts in signal transduction by specifically processing (a) proprotein(s) *in planta*, or whether the enzyme has a broader substrate specificity and hydrolyses proteins it comes into contact with irrespective of their primary structure. There is as yet little conclusive evidence for an involvement of plant subtilases in specific proprotein or prohormone processing, although several instances have been reported. These include a report of the *in vitro* proteolytic processing of a membrane bound leucine-rich repeat (LRR) protein by the subtilase P69-A, which may mediate recognition and interaction events in the extracellular matrix of tomato during pathogenesis, in particular (Tornero *et al.*, 1996a). A derivative of systemin (an inducer of protease inhibitors), was shown to be processed by a membrane bound tomato protease, SBP50, which resembles the known subtilase prohormone convertase Kex2 (Schaller and Ryan, 1994).

It was proposed to conduct preliminary work into whether the protease cleaves specific protein substates or a broad range of substrates. The consensus so far is that unlike mammalian subtilases, the plant subtilases seem to have very little in the way of substrate specificity, being capable of hydrolysing most proteins. The substrate specificity of newly purified proteases, including the plant subtilisin-like proteases, have commonly been studied using chromo- and fluorogenic peptide substrates and the oxidized B-chain of insulin. It was proposed to use these substrates, as well as native proteins, to explore the activity of Ara12 protease.

A variety of compounds are known to act as protease inhibitors and a number of these inhibitors were studied for their ability to abrogate the protease activity of Ara12. Known inhibitors of serine proteases were examined, as well as inhibitors of cysteine and acidic aspartic proteases. This addresses the fundamental mode of catalysis of the enzyme, which has been predicted to be a subtilisin-like serine protease. It may even lead to approaches with which the role of Ara12 could be specifically targeted for research. Inhibitors could also potentially form the basis of a normal method of controlling these enzymes *in planta*, particularly in the light of the discovery of a number of subtilisin inhibitors in plants.

6.2 The pH profile of Ara12 protease activity

Enzymes tend to be stable over a limited pH range, outside of which the tertiary structure of the protein can be modified as charges on ionizable amino acid residues change. This can lead to denaturation of the enzyme. The effect of pH on Ara12 protease activity was investigated using the purified enzyme. The pH found in the primary cell wall of plants can vary considerably and is particularly low during acid growth associated with wall extension (McQueen-Mason and Cosgrove, 1995).

The pH range over which the buffering action of different buffers is effective varies greatly. Therefore different buffers were used over the pH range tested to make allowance for this. Sodium acetate was used between pH4.0-6.5 and Tris was used between pH6.5-10.0. Protease assays using FTC-casein were carried out in a total volume of 50 µl. The FTC-casein substrate (1% w/v) was made up in autoclaved Milli Q water. The protease activity of Ara12 was measured in 50 mM NaOAc.HOAc solutions of varying pH (pH4.0,

4.5, 5.0, 5.5, 6.0 and 6.5) and in 50 mM Tris.HCl solutions of varying pH (pH6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5 and 10.0). The final assay mixture contained 10 µl of substrate and 38 µl of the appropriate buffer solution. Proteolysis was initiated by adding 2 µl of purified Ara12 enzyme solution. Samples were incubated at 37°C for 8 hours. Relative protease activities were determined and plotted against the relevant pH values (see Figure 6.1).

The pH optimum for the proteolytic activity of Ara12 protease was determined as approximately pH5.5. However the results show that Ara12 retains activity between pH5.0 and 10.0. Other plant subtilisin-like proteases have been shown to be active over a broad pH range, for example cucumisin has been shown to be stable between pH5.0-9.0 (Kaneda *et al.*, 1984).

6.3 Temperature dependence of Ara12 protease activity

In general enzyme activity increases with increasing temperature until temperatures are reached at which the enzyme becomes unstable and becomes irreversibly inactivated. This process occurs continuously, however increasing the temperature increases the rate of inactivation. The stability of an enzyme will also depend upon the length of time it is incubated at a certain temperature. The stability of Ara12 protease may be adversely affected by long incubations at high temperatures. For this reason it was particularly important to determine the temperature dependence of Ara12 protease activity over the relatively long periods of time (up to 18 hours) at which some of the assays were performed during the course of this work. This information was used to assess a suitable temperature at which to perform the assays.

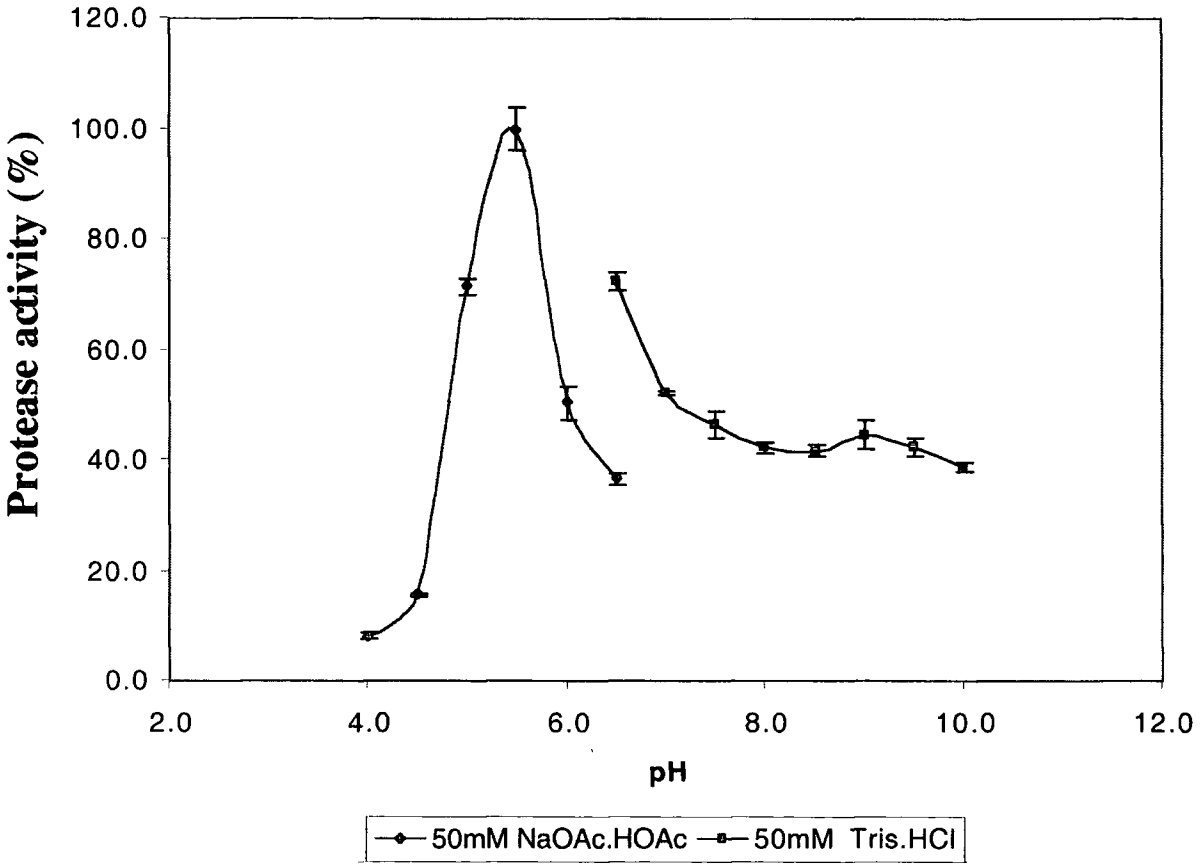


Figure 6.1 The effect of pH on Aral2 protease activity. Protease activity was measured using purified Aral2 protease in 50mM sodium acetate (NaOAc.HOAc) buffer from pH 4.0-6.5 and in 50mM Tris.HCl buffer from pH 6.5-10.0 with FTC-casein. Assay samples were incubated at 37°C for 8 hours. The mean and variance of data from duplicate experiments are shown.

The purified Ara12 protease was used to monitor the dependency of its activity upon the temperature of the reaction. Protease activity was measured with the FTC-casein assay, as described in Chapter 2. Samples were pre-incubated for 10 minutes at a range of different temperatures (8°C, 13°C, 20°C, 30°C, 37°C, 42°C, 50°C and 60°C) prior to addition of Ara12 protease. The samples were incubated for 18 hours at the temperatures indicated.

The relative activities were plotted against the temperature values (see Figure 6.2). Relative activity increased with increasing temperature up to a maximum temperature for the Ara12 induced proteolysis, which appeared to be around 42°C. At higher temperatures the relative activity decreased. At 50°C approximately 75% of the activity found at 42°C was observed. At 60°C that figure had dropped to 60%. At these temperatures the Ara12 protease may not be stable for long periods of time. Although the maximum activity in these assays over 18 hours was observed at 42°C most assays were performed at 37°C, which showed 90% of the activity seen at the maximum.

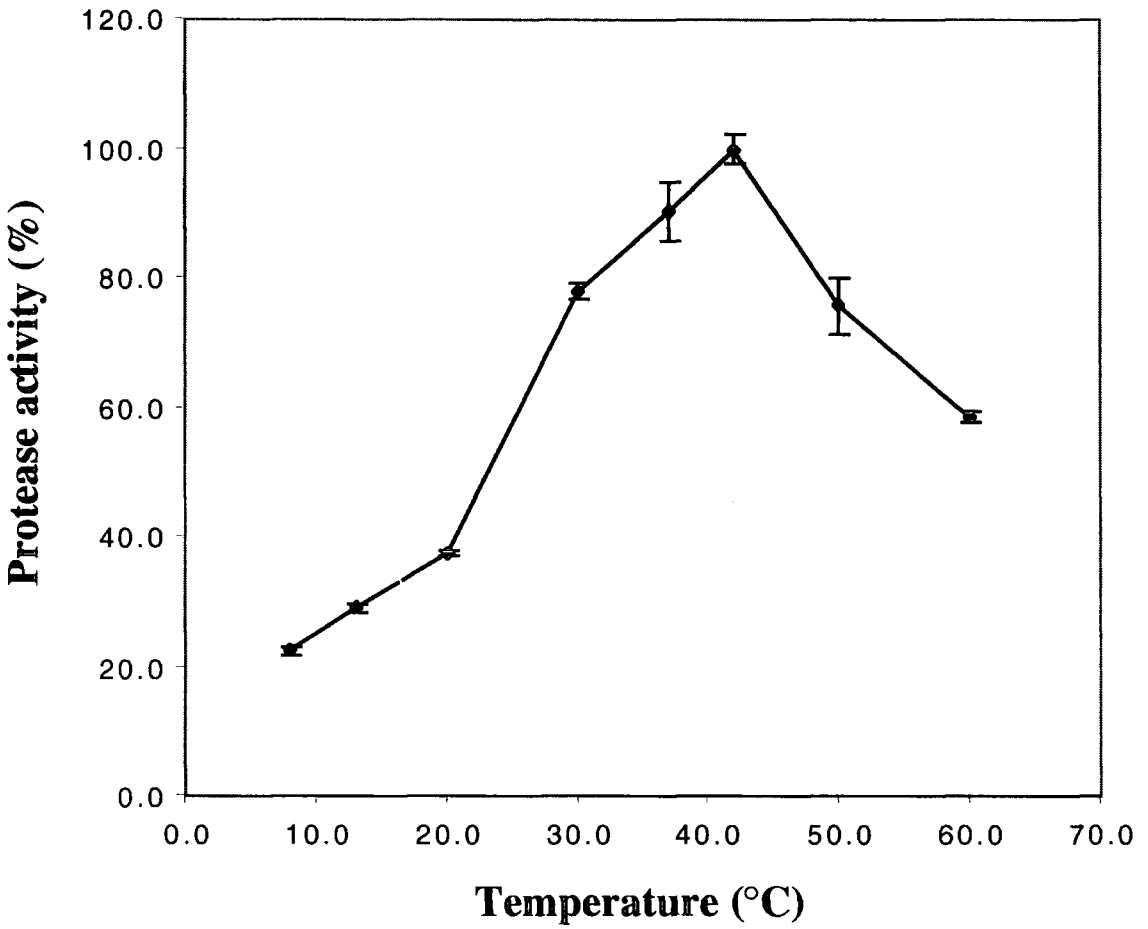


Figure 6.2 The effect of temperature on purified Ara12 protease activity. FTC-casein substrate was pre-incubated at the temperatures indicated (between 8-60°C) for 10 minutes before initiating substrate hydrolysis by adding purified Ara12 protease. The reaction was terminated after 18 hours by addition of TCA to a final concentration of 3.5% and the protease activities were determined. Mean relative protease activities and variance of data from duplicate experiments are shown.

6.4 The effect of cofactors on Ara12 protease activity

A cofactor is generally defined as a non-protein substance that forms a complex with certain enzymes and can be essential for their activity. Cofactors may be metal ions or coenzymes. The proteolytic activity of subtilisin from prokaryotes is known to be activated in the presence of calcium ions (Bond, 1996).

The activity of purified Ara12 subtilisin-like protease was examined to establish if activation of the enzyme could be brought about in the presence of increased calcium or magnesium ions. In addition the effect of EDTA and EGTA salts upon Ara12 activity was studied. These compounds act as chelating agents, effectively removing cations, such as Ca^{2+} and Mg^{2+}, from circulation.

Each 50 µl assay contained a final concentration of 0.2% FTC-casein and 2 µl (120 ng) of purified protease solution. Assay mixtures contained varying amounts of $CaCl_2$ or $MgCl_2$ (0-10 mM final concentration) in Tris.HCl, pH7.5. Relative protease activities were determined and plotted against the relevant salt concentrations (see Figure 6.3). An apparent activation of Ara12 activity was seen in the presence of calcium chloride. At 1 mM $CaCl_2$ the activity of Ara12 was about 145% of that with no added salt. As more $CaCl_2$ was added this figure rose. A stimulation of activity of around 60% was observed at 10 mM $CaCl_2$ above that found when no salt was added (ie. 160% of the no salt equivalent). Magnesium ions did not cause such a large activation of activity, however at a final concentration of 10 mM $MgCl_2$, approximately a 20% stimulation of activity was seen. Surprisingly, the chelating agents tested, EDTA and EGTA, caused only a very slight decrease in protease activity at final concentrations of 10 mM. Samples containing either

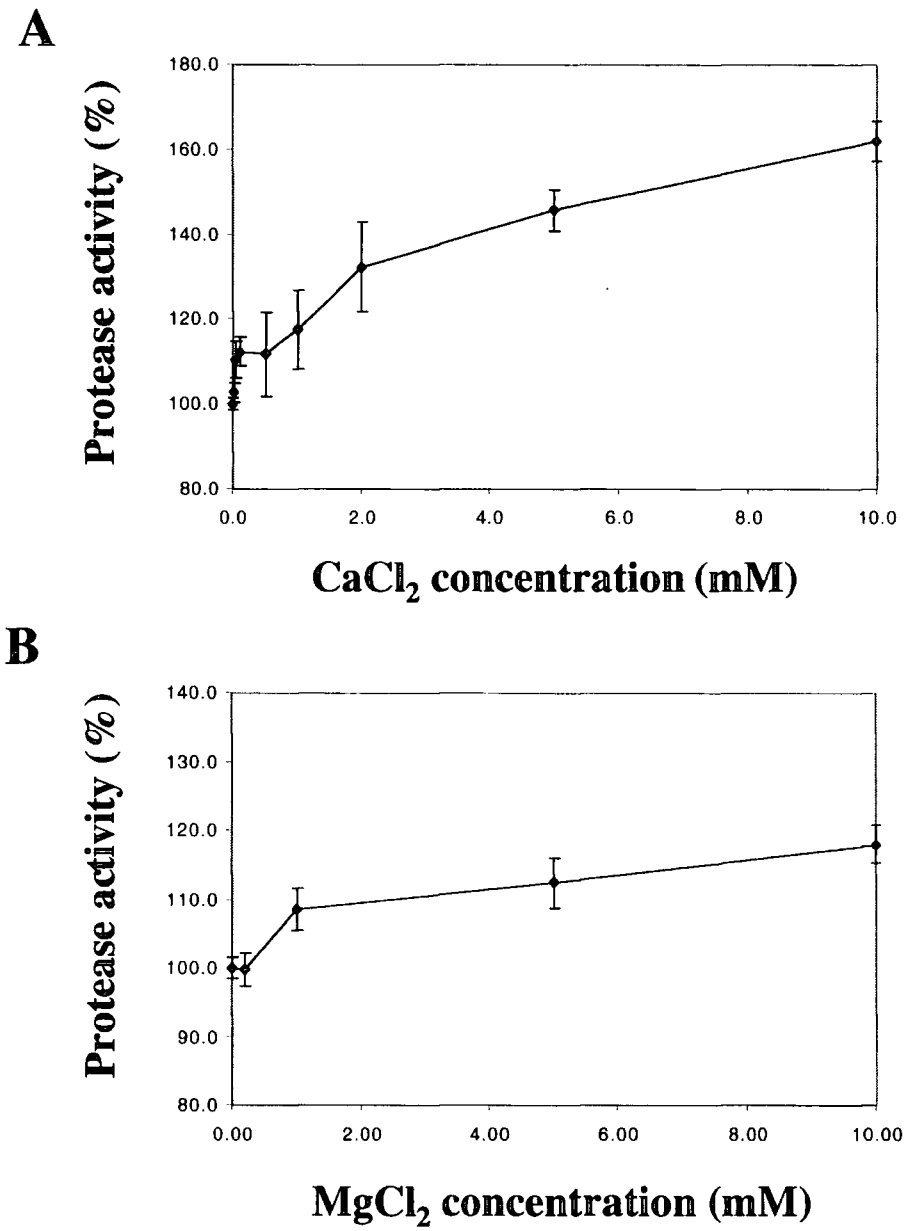


Figure 6.3 The effect of $CaCl_2$ and $MgCl_2$ on purified Ara12 protease activity. Protease activities were measured with FTC-casein in the presence of differing concentrations of $CaCl_2$ (**A**) or $MgCl_2$ (**B**) and are shown as percentage activity relative to equivalent samples with no salt added. The mean and variance of data from duplicate experiments has been shown.

10mM EDTA or EGTA showed 98% or 93% of the control activity, respectively (see Table 6.1). This might indicate that Ara12 protease may be stabilized, but is not necessary dependent on calcium ions.

6.5 Studies into the substrate specificity of Ara12 protease

Purified Ara12 protease was used to investigate the cleavage of native proteins, oxidized insulin B chain and artificial peptides to characterize the substrate specificity of the enzyme.

6.5.1 Cleavage of native proteins

Having established that the Ara12 protein, which was predicted to be a subtilisin-like protease, displayed proteolytic activity towards a fluorogenically labelled casein substrate (see Chapter 5), it was decided to determine whether the protease showed a similar activity to other commercially available native proteins, before examining native proteins which were likely to be in close physical contact with Ara12: namely other *Arabidopsis* cell wall proteins.

2 μg of the purified native proteins bovine carbonic anhydrase, bovine serum albumin and ovalbumin and 250 ng of bovine κ-casein (all purchased from Sigma) were incubated with and without 250 ng of purified Ara12 subtilisin-like protease in 20 mM Tris.HCl, pH7.5, 10 mM $CaCl_2$ at 37°C for 48 h. The total volume of each sample was 40 μl. Samples were then compared by SDS-PAGE analysis on a 10% polyacrylamide gel (see

Figure 6.4). It is usual to use high enzyme:substrate molar ratios for the cleavage of native proteins. This is because the proteins retain their folded structure and their disulphide bridges and therefore only part of the protein is available for the protease to cleave. Therefore enzyme:substrate w/w ratios of 1:8 and 1:1 were used in this investigation. A long incubation time was used, as is also common, to allow the protease to hydrolyze the substrate protein.

The proteins examined were all degraded by the protease solution added, as can be seen in Figure 6.4. A number of discrete bands, of a size smaller than that of the substrate molecule, were visible on the gel in lanes 6 and 8. These were presumably proteolytic degradation products of BSA and ovalbumin respectively. Therefore, to a certain extent, Ara12 protease was cutting these proteins into polypeptides of a distinct size under the conditions described here. But this effect was probably caused purely by the availability of the substrate peptide linkages, rather than necessarily revealing a fundamental substrate selectivity of the protease.

Cell wall protein extracts generated from *Arabidopsis* suspension cultured cells were used to investigate the effect Ara12 protease would have on them. These proteins are likely to be in close contact with Ara12 in the plant extracellular matrix and it was deemed to be of particular interest to find out how these proteins would be affected *in vitro*. This may give some idea of what the situation is like under the specific circumstances of pH and temperature experienced *in planta*.

Arabidopsis suspension culture cells were harvested 5 days after subculture. Cell wall material was prepared from these cells using a French press and layering the homogenised

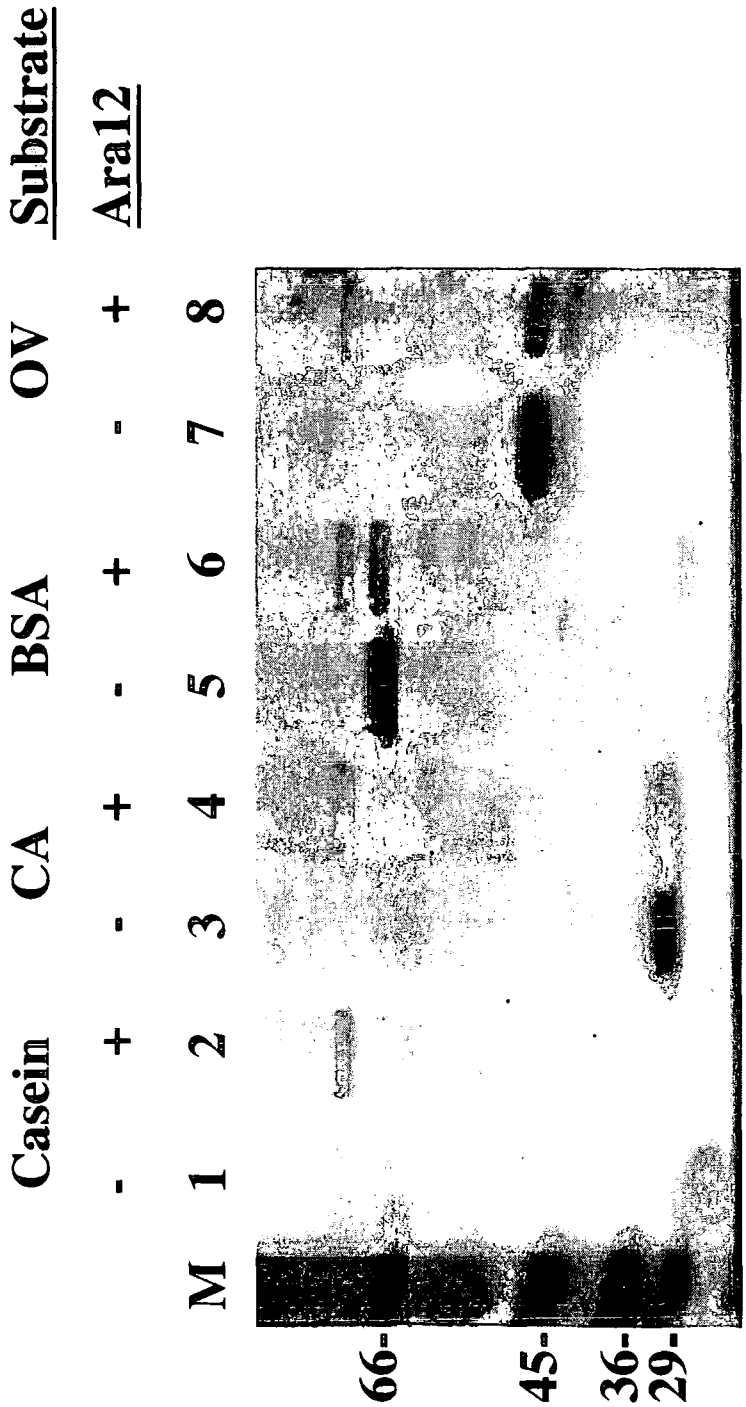


Figure 6.4 Digestion of proteins with purified Ara12 protease. 0.25μg of bovine kappa casein and 2μg bovine carbonic anhydrase (CA), bovine serum albumin (BSA) and ovalbumin (OV) were incubated with or without 0.25μg Ara12 protease at 37°C for 48 hours. The samples were analysed by SDS-PAGE on a 10% polyacrylamide gel. All the four proteins tested were degraded by the protease. Molecular masses of the protein markers (M) are given in kDa.

cells onto 10% (v/v) glycerol. Cell walls sedimented to the bottom of the glycerol solution and were extracted with ice cold 0.2 M $CaCl_2$. The preparation of pure cell walls and cell wall extracts is described in detail in Chapter 2. 5 μg of the two *Arabidopsis* cell wall protein extracts were incubated with and without purified Ara12 protease in extract:enzyme w/w ratios of 50:1 and 25:1 in 20 mM Tris.HCl, pH7.5 at 37°C for 18 h.

Samples were analysed by SDS-PAGE on a 10% polyacrylamide gel (see Figure 6.5). Compared to the control lacking protease, shown in lane 1, no degradation of proteins in the cell wall preparation was detected in the presence of Ara12 protease, (see lanes 3 and 5) as the protein profiles appeared to remain the same. However, it was apparent that proteins in the $CaCl_2$ cell wall extract sample were progressively degraded in the presence of increasing amounts of Ara12 protease. In Figure 6.5, lane 2 shows the $CaCl_2$ cell wall extract with no added protease. In the presence of 100 ng or 200 ng of Ara12 protease, shown in lanes 4 and 6 respectively, the protein profile observed demonstrated that proteolysis of the proteins occurred.

The reason that protein hydrolysis of $CaCl_2$ extracts of the cell wall was seen, whilst no significant hydrolysis of the original cell wall protein preparation could be seen, could be attributed to the presence of $CaCl_2$ used in the extraction process. Ca^{2+} ions have been shown to act as activators of subtilisin and have been shown in this work to stimulate activity of Ara12 protease, explaining why the proteins appear to have been preferentially degraded in the presence of Ca^{2+} ions. Degradation of native extracellular proteins would be consistent with this protease possessing a broad substrate specificity.

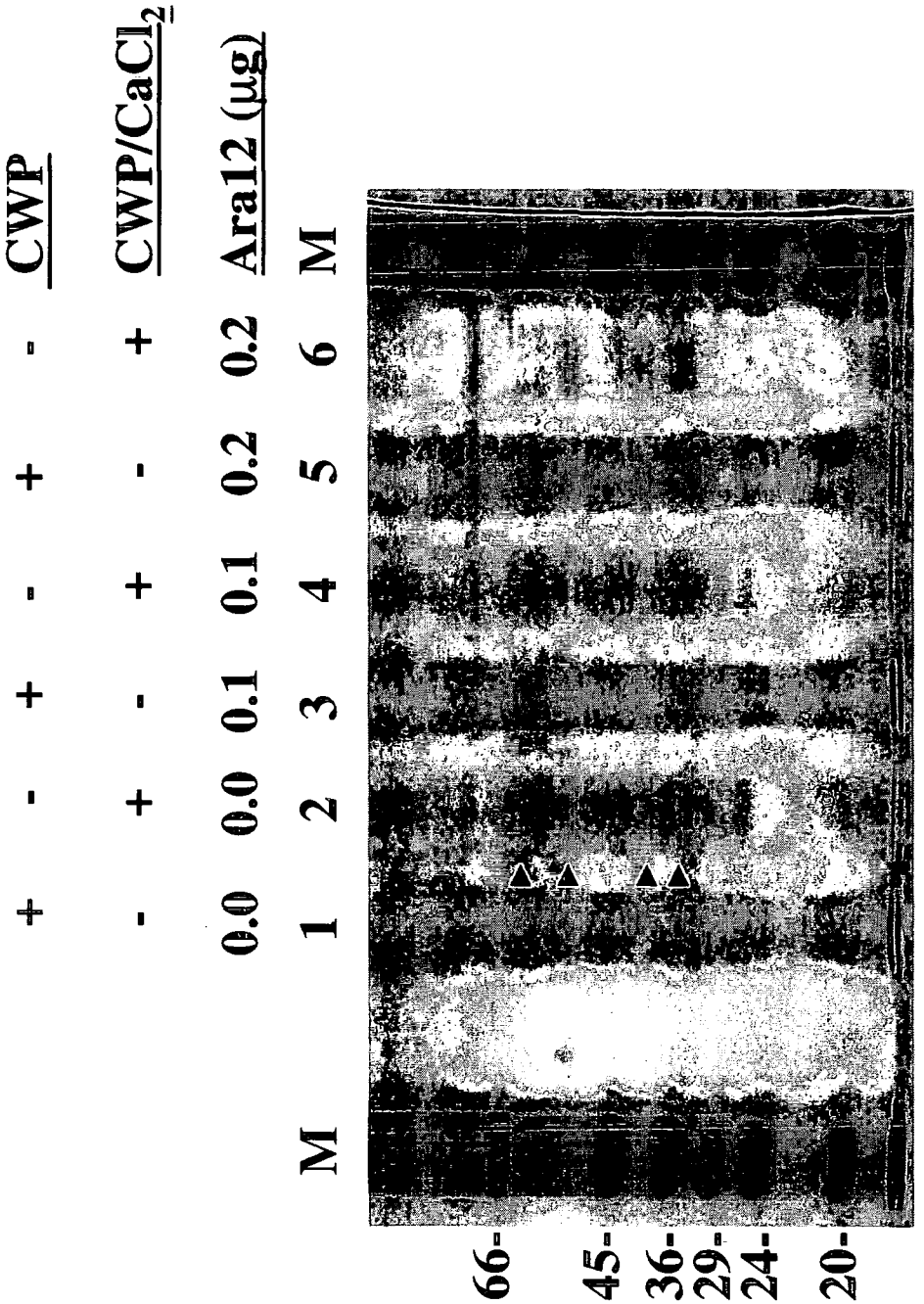


Figure 6.5 Digestion of *Arabidopsis* cell wall protein extracts with purified Ara12 protease. Cell wall protein (CWP) was prepared from growing *Arabidopsis* suspension cultured cells 5 days after subculturing. CWP (5µg) and a $CaCl_2$ extract of CWP (5µg) was incubated at 37°C for 18 hours with 0, 0.1 and 0.2µg of purified Ara12 protease. The protein samples were resolved on a 10% polyacrylamide gel. Molecular masses of the markers (M) have been given in kDa. The presence or absence of 5µg of CWP or a $CaCl_2$ extract of CWP has been shown by + or – signs. Where included Ara12 protease was added to a w/w ratio of protein extract: protease of 50:1 (0.1µg) or 25:1 (0.2µg). Proteins which appear to be degraded by Ara12 are indicated with arrowheads.

6.5.2 Cleavage of artificial peptides and oxidized insulin B-chain

Artificial substrates were used to determine whether Ara12 subtilisin-like protease would degrade certain peptide substrates, which other subtilisin-like proteases are known to cleave. One of the substrates used was Z-Gly-Gly-Leu-AMC, which is known to be cleaved by subtilisin Carlsberg (Kanaoka *et al.*, 1985). A second substrate, Boc-Leu-Arg-Arg-AMC.HCl, has been used to monitor cathepsin A and Kex2 activity (Mizuno *et al.*, 1987). Both substrates were purchased from BACHEM (UK) Ltd.

Purified Ara12 protease (1 μg) was incubated with 0.1 mM solutions of these substrates in 20 mM Tris.HCl, pH7.0, 10 mM $CaCl_2$ at 25°C. The Ex355 Em460 fluorescence measurement was followed with a fluorimeter over a 2 h period. Similarly, 1 μg of subtilisin Carlsberg (purchased from Sigma) was also used to monitor the degradation of these tripeptide substrates under identical conditions, for comparison.

It was found that the rate of hydrolysis of both of these substrates by purified Ara12 subtilisin-like protease was negligible, however as expected subtilisin Carlsberg caused hydrolysis of the Z-Gly-Gly-Leu-AMC substrate (data not shown).

When a novel protease has been isolated another standard way of exploring the specificity of the newly discovered activity has been to digest oxidized insulin B-chain, analyse the peptides released (normally by RP-HPLC) and compare them against the standard peptide maps (Wang and Carpenter, 1967; Butler *et al.*, 1987). Although in principle any denatured peptide or protein could be used as a substrate, insulin B-chain has traditionally been used by researchers, and therefore the new activity can readily be compared to established protease activities (Dunn, 1996).

Oxidized insulin B-chain (250 μg) was digested in 50 mM ammonium acetate, pH5.6, 2.5 mM $CaCl_2$ in a final volume of 250 μl with purified Ara12 protease (using a 50:1 w/w ratio of substrate:protease). The digestion mixture was incubated at 37°C and 50 μl samples were removed, added to an equal volume of 0.2% TFA and frozen in liquid nitrogen at 0, 2, 4, 8 and 24 h after addition of the enzyme. Samples were kept at -80°C until the peptide content was analysed by reverse phase HPLC (RP-HPLC). It was not possible to resolve the sample proteins or peptides, as the column used for this, a C8 2.1/3 RP-HPLC column, appears to have been faulty. Another way of generating a cleavage map for the Ara12 protease using small proteins is to use MALDI-TOF (matrix assisted laser desorption ionisation – time of flight) mass spectroscopy however this was not completed during the course of this work.

6.6 Inhibition of Ara12 protease activity

Various different protease inhibitors were incubated with purified Ara12 protease to assess their ability to inhibit this serine protease. The following protease inhibitors were tested at the final concentrations shown: PMSF (0.5, 1, 2 and 5 mM), AEBSF (0.5, 1 and 2 mM), DFP (0.1, 0.5 and 1 mM), benzamidine (1 and 5 mM), iodoacetamide (1 and 5mM), leupeptin (0.02 and 0.1 mM), pepstatin A (1 mM), and soybean trypsin inhibitor (0.04 and 0.4 mM). After incubation of purified Ara12 with the inhibitors for 1 hour at the final concentrations indicated at 20°C at pH7.5, residual activity was measured with FTC-casein. Reactions were initiated by addition of the substrate. The assay was performed at 37°C for 8h. Inhibitor solutions were freshly prepared immediately before use. PMSF was

prepared in isopropanol, DFP and pepstatin A were prepared in DMSO and all other inhibitors were prepared in autoclaved Milli Q water.

A significant reduction in activity was found with the serine protease inhibitors PMSF, AEBSF, DFP and soybean trypsin inhibitor compared to a control sample lacking inhibitor. See Table 6.1 and Figure 6.6 for the results of the inhibition studies. PMSF and its analogue AEBSF inhibited Ara12 protease, such that at the 2 mM level 55% and 43% of the activity remained, respectively. When the final concentration of PMSF was raised to 5 mM, approximately just 17% of the uninhibited proteolytic activity was detected. Only less than half the activity was found in the sample containing 1 mM DFP. Soybean trypsin inhibitor (SBTI), at a final concentration of 0.4 mM, also caused a large decrease in activity, with 70% of activity remaining. No inhibition was seen with 1 mM or 5 mM benzamidine, or with 1 mM or 5 mM iodoacetamide, and only a minor inhibitory action was registered with 0.1 mM leupeptin and 1 mM pepstatin A as approximately 95% of activity was retained in these samples. In addition to inhibitors, the reducing agent dithiothreitol (DTT) was used in assays to assess its ability to inhibit the proteolytic activity of Ara12 protease. Inclusion of DTT at final concentrations of 1 mM and 5 mM caused a 2-fold and 2.5-fold stimulation of activity, respectively. The presence of this reducing agent in the protease assays would be expected to maintain disulphide bridges formed between cysteine residues found in Ara12. This may lead to greater stability of the enzyme and may account for the increase in activity observed.

PMSF, AEBSF, DFP, benzamidine and soybean trypsin inhibitor (SBTI) act as inhibitors of serine proteases, and all of these inhibitors apart from benzamidine reduced the activity

Reagent	Final concentration (mM)	Activity (%)
PMSF	0.5	83.7
	1	71.3
	2	54.8
	5	16.6
AEBSF	0.5	66.2
	1	55.6
	2	42.8
DFP	0.1	92.6
	0.5	77.3
	1	47.5
Benzamidine	1	105.2
	5	102.6
Iodoacetamide	1	100.7
	5	109.6
Leupeptin	0.02	98.0
	0.1	95.2
Pepstatin A	1	95.7
Soybean trypsin Inhibitor (SBTI)	0.04	85.7
	0.4	70.2
Dithiothreitol (DTT)	1	206.4
	5	251.7
EDTA	1	99.9
	10	97.9
EGTA	1	102.7
	10	93.1
Ca^{2+}	5	145.8
	10	162.1
Mg^{2+}	5	112.5
	10	118.1

Table 6.1 Effect of inhibitors and activators on the activity of purified Ara12 subtilisin-like protease. The inhibitors were pre-incubated with purified Ara12 protease for 1 h prior to addition of substrate. For EDTA, EGTA, $CaCl_2$ and $MgCl_2$ the assays were performed without a pre-incubation step and these results are the means of duplicate experiments.

Figure 6.6 The effect of protease inhibitors on Ara12 subtilisin-like protease activity. Purified Ara12 protease (150ng per assay) was incubated in the presence of the following known protease inhibitors: PMSF, AEBSF, DFP, benzamidine, iodoacetamide, leupeptin, Pepstatin A and soybean trypsin inhibitor (SBTI) at the final concentrations given for 1 hour at 20°C. Residual protease activity was monitored using 0.2% (w/v) FTC-casein as the assay substrate. Protease activities have been shown as percentages relative to controls which lacked the inhibitor in the pre-incubation mixture.

The following inhibitors were tested at the concentrations indicated:
0.5mM PMSF
1mM PMSF
2mM PMSF
5mM PMSF
0.5mM AEBSF
1mM AEBSF
2mM AEBSF
0.1mM DFP
0.5mM DFP
1mM DFP
1mM Benzamidine
5mM Benzamidine
1mM Iodoacetamide
5mM Iodoacetamide
0.02mM Leupeptin
0.1mM Leupeptin
1mM Pepstatin A
0.04mM Soybean trypsin inhibitor (SBTI)
0.4mM Soybean trypsin inhibitor (SBTI)

Protease activity (%)
120
100
80
60
40
20
0
Inhibitor added (mM)
none
PMSF 0.5mM
1mM
2mM
5mM
AEBSF 0.5mM
1mM
2mM
DFP 0.1mM
0.5mM
1mM
Benzamidine 1mM
5mM
Iodoacetamide 1mM
5mM
Leupeptin 0.02mM
0.1mM
Pepstatin A 1mM
SBTI 0.04mM
0.4mM

of Ara12 subtilisin-like protease. Leupeptin and PMSF can act as inhibitors of cysteine proteases as well as serine proteases. In fact, leupeptin tends to reversibly inhibit trypsin-like enzymes. Leupeptin appeared to have a very slight inhibitory effect on Ara12 protease at the 0.1 mM level. Iodoacetamide, however, which acts as an inhibitor of cysteine proteases, showed no significant reduction in Ara12 activity. Pepstatin A acts as an inhibitor of acidic aspartic proteases and was also ineffective at inhibition of Ara12. Chelating agents, such as EDTA and EGTA can be used to reversibly inhibit metallo-proteases and may destabilise some serine proteases. From previous findings discussed in this chapter, it was discovered that EDTA and EGTA caused only a small reduction in protease activity at final concentrations of 10 mM. Inhibition studies confirm Ara12 protease as a serine protease and identify DFP, soybean trypsin inhibitor, PMSF and AEBSF as effective inhibitors of its activity. The serine protease inhibitor benzamidine was found to be ineffective at Ara12 inhibition.

6.7 Conclusions

The Ara12 protease has been shown to have a broad pH range over which its activity was maintained. The enzyme appeared to be stable between pH5 and pH10 and showed an optimum of approximately pH5.5. Other plant subtilisin-like proteases isolated also display high activities over a broad pH range. Cucumisin, a subtilisin-like protease isolated from melon mesocarp, has a similar broad pH range of 5.0-9.0, with a rather high optimum pH for activity of pH10.0 (Kaneda *et al.*, 1984). Taraxalisin, a subtilisin-like protease isolated from dandelion roots, has a pH range of between 6.0-9.0 and an optimum of pH8.0

(Rudenskaya *et al.*, 1998), whilst macluralisin, a subtilisin-like protease isolated from *Maclura pomifera* fruits, has been reported to have a pH range of between 7.0-9.0 with an optimum pH of 8.5 (Rudenskaya *et al.*, 1995). In general plant subtilisin-like proteases have neutral or alkaline pH optima. Recently a tomato subtilisin-like protease, LeSBT1, has also been shown to have an acidic pH optimum (Janzik *et al.*, 2000). It should perhaps be borne in mind that the culture medium into which the *Arabidopsis* cells were subcultured and maintained was somewhat acidic at pH5.7. The environment in which Ara12 was found *in planta*, the extracellular matrix, would also be expected to be an acidic one, and hence it may not be entirely surprising to find an acidic pH optimum for this enzyme.

Plant subtilisin-like proteases are generally rather thermostable, with the exception of taraxalisin (Bogacheva, 1999) and Ara12 protease showed approximately 60% activity at 60°C with maximal activity estimated at above 40°C under the conditions described. At lower temperatures, 20°C and under, the activity dropped to levels below 40% of the maximum.

Ca^{2+} ions appeared to activate Ara12 protease activity, with a 60% stimulation of activity at a concentration of 10 mM $CaCl_2$ above a control lacking added salt. Magnesium ions did not cause such a large activation of activity, however at a final concentration of 10 mM $MgCl_2$, approximately a 20% stimulation of activity was seen. As the chelating agents tested (at final concentrations of 10 mM) caused only a very slight decrease in protease activity it may indicate that Ara12 protease may be stabilized, but is not necessary dependent on calcium ions.

Known serine protease inhibitors, PMSF, AEBSF and DFP were demonstrated as having an inhibitory effect on Ara12 protease activity. DFP phosphorylates the active site serine residue (Ho and Hoskins, 1983), PMSF forms a covalent complex with serine or histidine residues of the active site (James, 1978), whilst AEBSF acylates the active site of serine proteases (Rich *et al.*, 1984). Soybean trypsin inhibitor, another serine protease inhibitor, also decreased the catalytic capability of Ara12. Other inhibitors such as the cysteine protease inhibitors leupeptin and benzamidine had no significant effect on the activity of this protease. Presumably benzamidine's inability to inhibit Ara12 goes some way towards explaining the inability of the protease to bind to benzamidine Sepharose, used in affinity purification of some serine proteases. Pepstatin A, an inhibitor of aspartic proteases also had very little effect on Ara12 activity. Taken together these results were largely predictable given the status of subtilisin-like proteases as serine proteases.

Substrate specificity was investigated using pure preparations of known native proteins, such as BSA, ovalbumin and bovine carbonic anhydrase. Degradation of all native proteins tested was witnessed over a 48 hour period. This indicated that Ara12 protease is unlikely to possess the very limited substrate specificity that would be expected for a protease responsible for processing one or several specific proproteins. What seems more appropriate given this evidence is that Ara12 shares features with other plant subtilases, in having a broad substrate specificity. Ordinarily plant subtilases hydrolyse peptide bonds formed by the hydrophobic amino acids, leucine, phenylalanine and tyrosine (Bogacheva, 1999), although they show specificity for a range of other bonds. Notably their specificity is wider than that shown by the bacterial enzyme subtilisin BPN'.

Cell wall proteins extracted with $CaCl_2$ were digested with Ara12 protease and it seemed that all the proteins in the sample were degraded by the enzyme. Proteins expected in the extract include cellulases, glycoproteins, expansins and xyloglucan endoglycosylase (Robertson *et al.*, 1997). Important structural proteins, such as extensins, are found in typical plant primary cell walls. Cleavage of these proteins at selected sites and times in the plant's lifecycle could help to promote plant expansion growth. Growth is restricted by the rather rigid architectural framework imposed mostly by the cellulose-hemicellulose, pectin and extensin protein networks. The involvement of expansins in allowing growth by wall-loosening is widely accepted, however the extent to which polysaccharide hydrolases, for example, endoglucanases and pectinases, or xyloglucan endoglycosylases are involved is more uncertain (Cosgrove, 1999). Even less is known about a possible loosening of the extensin network by proteolytic enzymes and much work is needed in this field in general to understand expansion growth in detail.

Artificial peptide substrates attached to fluorogenic or chromogenic compounds were used to investigate the substrate selectivity. The substrates used were, firstly, a typical substrate used to monitor subtilisin activity, namely Z-Gly-Gly-Leu-AMC. No hydrolytic activity was detected against this substrate with Ara12 however, as expected subtilisin Carlsberg degraded the substrate. No hydrolytic activity was detected against the second artificial substrate used, Boc-Leu-Arg-Arg-AMC.HCl, using either Ara12 protease or subtilisin Carlsberg. This substrate has been used to monitor Kex2 and other dibasic processing enzymes, characteristically involved in proprotein and prohormone processing (Barr, 1991).

Chapter 7

General discussion

7.1 General discussion

A large number of subtilisin-like proteases appear to be present in the model plant *Arabidopsis thaliana*, judging by the results of database searches and Southern blot analysis. The study of the phylogenetic relationships between proteins can reveal a great deal about the function of those proteins, as has been shown with the protein kinases (Hanks *et al.*, 1988). The dendrogram in Figure 7.1 shows the phylogenetic relationship of the known higher plant subtilisin-like proteases. Some ideas concerning the putative functions of these proteases will be put forward here, formed purely on the basis of sequence homologies between related plant proteases. Slpg and Slpj, for example, have closest homology to tmp and LIM9 proteases, which are both known to be involved in microsporogenesis (Riggs and Horsch, 1995; Taylor *et al.*, 1997). The latter two enzymes have actually been placed together in the tree on account of their shared sequence conservation, but it transpires that they also play similar roles in plant development. This alone indicates that there is predictive value in the examination of phylogenetic relationships amongst proteins, as has been illustrated in the past (Hanks *et al.*, 1988). It remains to be seen whether Slpg and Slpj are involved in microsporogenesis or some other aspect of plant development. It is still somewhat unclear, but LIM9 may have a role in tapetal cell apoptosis as microsporocytes reach the tetrad stage (discussed by Taylor *et al.*, 1997).

The tomato P69 proteases seem to be in a separate subfamily, being extremely closely related to each other. They are thought to be involved in pathogenesis (Jordá *et al.*, 1999). The *P69A* and *P69B* genes are highly expressed after infection of tomato plants with citrus exocortis viroid (Tornero *et al.*, 1996b; Tornero *et al.*, 1997). *P69B* and *P69C* show increased expression following infection with *Pseudomonas syringae* and are also upregulated by salicylic acid.

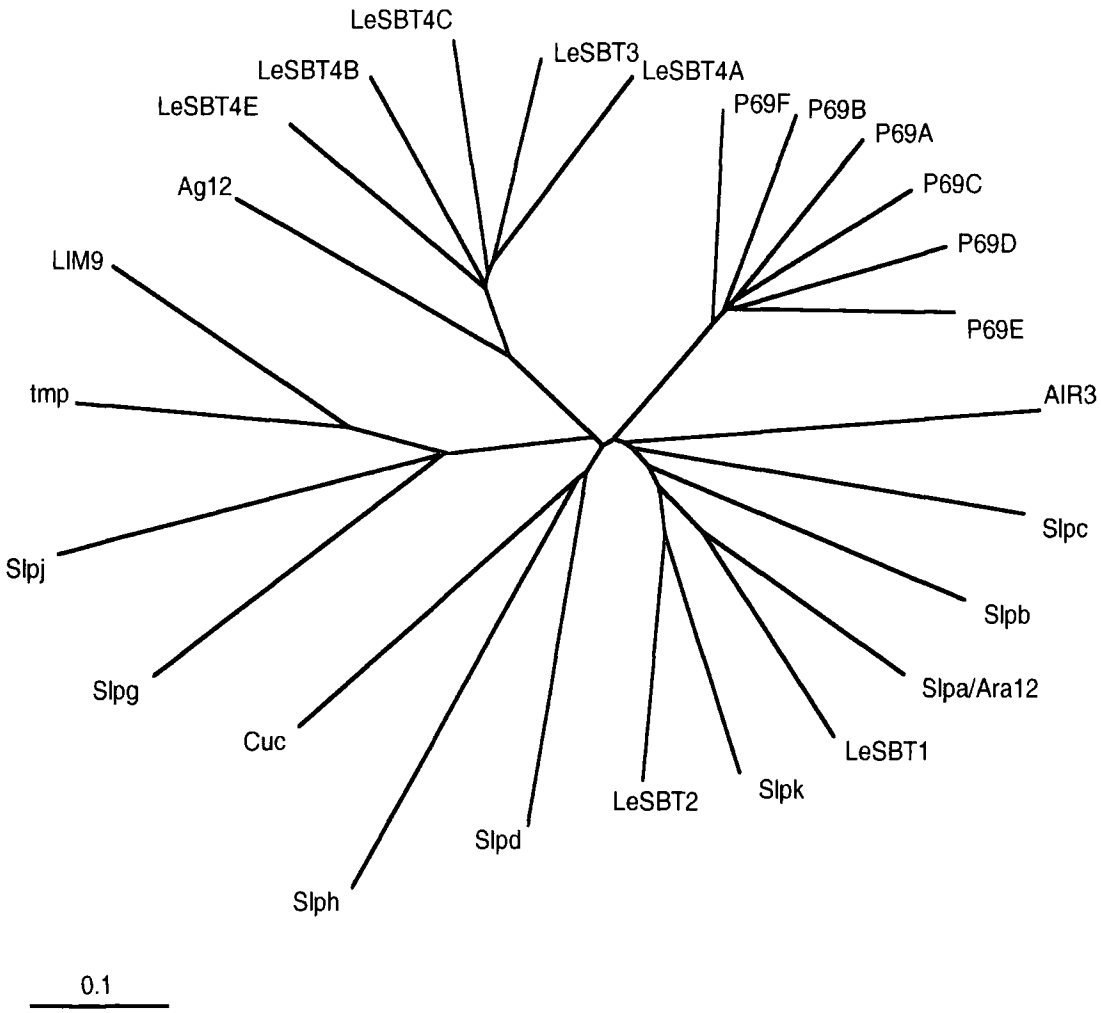


Figure 7.1 Phylogenetic relationship of higher plant **s**ubtilisin-like **p**roteases (**Slps**). Amino acid sequences were aligned using the ClustalW1.7 program and the unrooted phylogenetic tree was viewed using TreeView1.5.3. The amino acid sequences of the following proteins shown above have been published: cucumisin from melon (**Cuc**; Yamagata *et al.*,1994), **Ag12** from alder and **Ara12 or Slpa** from *Arabidopsis* (Ribeiro *et al.*, 1995), **LIM9** from lily (Taylor *et al.*, 1997), **AIR3** from *Arabidopsis* (Neuteboom *et al.*, 1999) and **tmp** (Riggs and Horsch, 1995), **P69A** (Tornero *et al.*, 1996b), **P69B** (Tornero *et al.*, 1997), **P69C-F**, **LeSBT1-3** and **LeSBT4A-E** (Meichtry *et al.*, 1999), all from tomato. The **LeSBT4D** sequence from tomato was not included here, because a large part of this cDNA has not been cloned or sequenced. The *Arabidopsis* amino acid sequences **Slpb**, **c**, **d**, **g**, **h**, **j** and **k** were obtained in searches of the NCBI database. Minimal distance between sequences is given in PAM (accepted point mutations per 100 residues).

Slpa, Slpb, Slpc, Slpk, LeSBT1, LeSBT2 and AIR3 appear to form another, less well defined, homology cluster. It is thought that AIR3 could be involved in lateral root formation (Neuteboom *et al.*, 1999). The XSP1 sequence has not been represented in Figure 7.1, but has highest homology with this group of proteases and is thought may be involved in xylem development (Zhao *et al.*, 2000). It can not be ruled out that the other members of this homology cluster have a direct or indirect role in plant development, in the tissues where expression has been demonstrated.

Cucumisin, Slpd and Slph form another branch of the unrooted phylogenetic tree. Cucumisin is found in large quantities in maturing melon fruit (Yamagata *et al.*, 1994), but it's precise function is unclear.

Ag12, LeSBT3 and LeSBT4 proteases could potentially form a different subfamily. Ag12 seems to be involved in the early stages of actinorhizal nodule development (Ribeiro *et al.*, 1995). Ag12 might be involved in processing a cell wall protein, or degrading all the cell wall proteins of infected cells, in order to establish a symbiotic relationship between alder and *Frankia* bacteria. How this relates to the poorly characterised LeSBT3 and LeSBT4 proteases is not clear.

Of course care should be taken not to draw too many hard and fast conclusions from these observations, but it certainly does spur a great deal of thought for future experimentation. Some of the most useful ways to study this would be to produce plants which have the gene of interest "knocked out" by generating antisense plants or selecting for plants containing protease genes disrupted by a T-DNA insertion. The phenotype of those plants could be examined under normal growth conditions and conditions of biotic and abiotic stress. Normal cellular and subcellular locations of the proteases could be determined in a variety of ways, as well as finding out what might alter protease levels in the plant.

It would be interesting to know whether the functions of plant subtilisin-like proteases have a common theme running through them. For example, are they all involved in specific targeted apoptotic events? Mammalian subtilisin-like proteases preferentially cause cleavage at dibasic sites. They appear to have an important role as preprotein convertases. A host of proteins are processed, at paired basic residues, from larger precursors, including peptide hormones, neuropeptides, growth factors, receptors and adhesion molecules (Nakayama, 1997). Prokaryotic and plant subtilisin-like proteases appear to lack this narrow substrate specificity *in vitro*. Evidence for a broad substrate specificity has come from biochemical analysis of cucumisin (Kaneda and Tominaga, 1975), macluralisin (Rudenskaya *et al.*, 1995) and Slpa/Ara12 protease and is discussed in other chapters. Rather than acting as preprotein or prohormone convertases, these, mainly extracellular, proteases might weaken or destroy proteins in the relatively rigid cell wall to allow the plant to undergo a developmental transition of one form or another. The stimulus for protease expression might be a developmental cue or infection by pathogens or symbiotic bacteria. Apoptosis is a part of normal development and cell death is involved in typical plant responses to pathogen attack. Alternatively Slps may act in protein processing as part of signal transduction pathways (Schaller and Ryan, 1994), or there may be members of this family operating both in signal transduction and in apoptosis or localised proteolytic degradation. It will be interesting to see whether the characterized Slps and the large number of uncharacterized ones act directly on their protein environment *en masse* or have specific protein substrates. No doubt, precisely because of the non-specific nature of the plant subtilisin-like proteases which have been studied, they will need to be under strict temporal and spatial regulation. Enzyme activities may also be regulated by specific inhibitors and/or other molecules, such as calcium salts.

Using DNA sequences obtained during the ongoing sequencing of the *Arabidopsis* genome, and held on public-access databases, oligonucleotide primers complementary to various putative subtilase genes were designed and synthesized. These primers were used to amplify DNA sequences by PCR predicted to code for *Arabidopsis* subtilases, including *slpb*, *slpc* and *slpk*. The tissue specificities of these genes have been examined by Northern blotting.

Another way of determining the sites and regulation of subtilase gene expression in *Arabidopsis* includes the study of transgenic GUS (β-glucuronidase) or GFP (green fluorescent protein) reporter gene fusion plants. Subtilase gene promoters can be ligated to these reporter genes and these DNA constructs can be used to transform plants. Subtilase gene expression can be monitored by observing any colour or fluorescence changes in the tissue associated with expression of the reporter genes. The effects of exposure of the plants to different hormones, environmental conditions or pathogens can be examined using this method.

Preliminary experiments using *Arabidopsis* cell suspension cultures have suggested that protease activity of the culture filtrate increases (as determined using an FTC-casein assay) in the presence of increasing amounts of both NAA (2-naphthaleneacetic acid) and IAA (indole-3-acetic acid). These results are not discussed in the results and are only mentioned briefly here, as the experiment was only performed once. Ara12 protease is known to constitute the dominant proteolytic activity in the filtrate of these cell cultures under normal incubation conditions, although another proteolytic activity has also been detected (the origin of which is unknown). It can be speculated that Ara12 protease activity is increased and/or more Ara12 protease is generated by the cells in response to elevated levels of auxin, but there is as yet no firm proof of this. It is thought that expression of another subtilase gene from *Arabidopsis*, *AIR3,* is

upregulated in response to increasing levels of auxin (Neuteboom *et al.*, 1999). Ara12 levels would need to be monitored using the Ara12 antibody to categorically determine whether addition of auxin can alter *ara12* expression in *Arabidopsis* suspension cultured cells.

Polyclonal antibodies have been generated which appear to recognize the Ara12 protein in crude protein extracts from *Arabidopsis* tissues. This has offered a method of determining the location of the Ara12 protease, rather than looking at where the *ara12* gene is transcribed, or where it is envisaged that it is transcribed. The Ara12 protein is located mostly in silique and stem tissue, but is present in all tissues examined. Furthermore, from immunocytological experiments with longitudinal sections of silique cells, Ara12 appears to be found in the apoplast.

The Ara12 subtilisin-like protease has been purified from *Arabidopsis* cell suspension culture filtrate by anion exchange and hydrophobic interaction chromatography. The protease showed relatively thermostable proteolytic activity against protein substrates, such as casein. Unusually, for a plant subtilase, its pH optimum was acidic, being approximately pH5.5 with acetate buffer. The proteolytic activity of Ara12 can be inhibited by the serine protease inhibitors DFP, PMSF and AEBSF and is stimulated in the presence of Ca^{2+} ions.

It would be interesting to see whether the enzyme is processed to a smaller activated protease at acidic pH, in a similar fashion to a homologue of Ara12: *Le*SBT1 from tomato (Janzik *et al.*, 2000). It would also be of interest to determine the precise specificity of this enzyme. This could be achieved by analysing the cleavage products resulting from the degradation of small protein substrates by Ara12 protease using reverse phase HPLC and protein sequencing or MALDI-TOF. Establishing the

substrate specificity would give further insight into the role of this pyrolysin (the family of subtilisin-like proteases to which Ara12 belongs) in the apoplast.

In general plant subtilases have been found to possess a fairly broad substrate specificity (Bogacheva, 1999). A broad substrate specificity has important implications for the function of this group of enzymes and would place them in roles more likely to involve targeted apoptotic events possibly in normal development (or "abnormal" developmental events such as pathogenesis or the establishment of symbiotic relationships) within plant tissues. Given the characteristics of these enzymes it makes them more likely to be involved in these kind of proteolytic events, rather than acting as specific preprotein convertases as is observed with mammalian subtilases.

Ara12 protease did not show proteolytic activity against the artificial substrate Boc-Leu-Arg-Arg-AMC.HCl, a substrate used to monitor proteolytic activity of Kex2 and other dibasic processing enzymes. This suggests that Ara12 may not be involved in proprotein processing, at least not in a way identical to characterized dibasic processing enzymes.

Other ways of investigating the role of the Ara12 protease include an antisense approach and the study of previously generated mutagenized plant lines (Feldmann, 1991). An *ara12* antisense construct has been generated and used to transform *Arabidopsis* plants. However no physiological differences were observed in these plants. This experiment is not discussed in the results presented in this work, but is briefly mentioned here. The cDNA coding for Ara12 proprotease was inserted between the *Kpn*I and *Xba*I sites of the plasmid pJO530 (a gift from Iain Graham from the University of Glasgow), based on the binary vector Bin19, in the antisense orientation. Flowering *Arabidopsis* plants were transformed with this construct by an

Agrobacterium-mediated vacuum infiltration method (Bent *et al.*, 1994). Seedlings from the transformed plants were selected according to their ability to grow on agar containing hygromycin (a gene conferring resistance to this antibiotic is found on the pJO530 plasmid). Normal seedlings display stunted root morphologies when grown under these conditions. Nine plants containing normal root systems and shown to contain the recombinant DNA were studied, however they did not appear to have abnormal morphologies. This may have been because Ara12 protease is not involved in the control of plant morphology. Alternatively, as there seem to be many subtilase enzymes in *Arabidopsis*, perhaps one of the homologues of Ara12 is capable of compensating for any reduction in active Ara12 protease.

A reverse genetics approach relying on previously generated mutant plant lines has also been explored. Thousands of transformants of *Arabidopsis* have been generated by Feldmann and co-workers by infecting germinating seed with *Agrobacterium tumefaciens* (Feldmann, 1991). The T-DNA (transfer DNA) of this bacterium is inserted into *Arabidopsis* genes and mutates them. Several unsuccessful attempts were made, during the course of this work, to identify a plant line which has a T-DNA insertion in the *ara12* gene, rendering the plants *ara12* mutants. It is possible that the plant DNA screened by PCR (from the different plant lines), did not contain any T-DNA insertions in the vicinity of the *ara12* gene, and therefore the plants would not be expected to be *ara12* "knockout" plants.

A different approach to examining the action of Ara12 may exist utilizing the polyclonal antibodies raised during this study. If these antibodies were added to growing *Arabidopsis* cell suspension cultures, the binding of antibodies to the protease may interfere with its proteolytic activity, offering the opportunity to study the role of this protease *in vivo*.

Clearly a great deal of work will be required to elucidate the functions of the Ara12 protease and all the other subtilisin-like proteases in *Arabidopsis* and the location and identities of their endogenous *in vivo* substrates. The use of mutants and antisense plants will undoubtedly be invaluable in this task. From our knowledge of plant subtilisin-like proteases so far, it appears that they may have a role in the modulation of plant morphogenesis during normal development and during pathogen attack. Hopefully in the future we will gain a better understanding of whether plant subtilisin-like proteases, such as Ara12, are involved in morphogenesis, and how the plant cell wall is modulated during normal development and as a reaction to biotic and abiotic stress. It remains to be seen to what extent the plant subtilisin-like proteases are responsible for the proteolytic processing of receptors, cell adhesion molecules or prohormones, or are involved in degrading most proteins (e.g. the network of extensins) they come into contact with. In either case these enzymes will need to be tightly regulated. For Ara12 protease part of this regulation may involve the pH of the compartment in which it performs its presumed proteolytic role. The pH of the apoplast, the apparent location of mature Ara12 protease, is reduced during elongation growth. An acidic pH is required for optimum hydrolytic activity of Ara12 protease. It could be hypothesized that Ara12 is involved in a weakening of the extensin network or cell adhesion molecules in the cell wall and intercellular spaces respectively, as part of a process of initiating localized plant cell growth. Ara12 is certainly found in tissue undergoing rapid cell growth: silique and stem tissue. In addition agents which are involved in wall relaxation may actually modulate developmental pathways. It has been shown for example, that local application of purified expansin protein to a living shoot apical meristem induces a leaf-like structure (Fleming *et al.*, 1999). Ara12 protease could alternatively be involved in satisfying the nutritional requirements of

plant cells during expansion growth. However, no substantial evidence exists suggesting a role for Ara12 in elongation growth or morphogenesis and many questions still remain unanswered.

References

Akiyama, Y. and Kato, K. (1982). An arabinoxyloglucan from extracellular polysaccharides of suspension-cultured tobacco cells. *Phytochemistry* **21**, 2112-2114.

Albersheim, P., An, J., Freshour, G., Fuller, M.S., Guillen, R., Hahn, M.G., Huang, J., O'Neill, M., Whitcombe, A., Williams, M.V., York, W.S. and Darvill, A. (1993). Structure and function studies of plant cell wall polysaccharides. *Biochem. Soc. Trans.* **22**, 374-378.

Alexander, C.M. and Werb, Z. (1991). Extracellular matrix degradation. In *Cell Biology of Extracellular Matrix*, (ed. Hay, E.D.), Plenum Press, New York, pp.255-302.

Anderson, E.D., Van Slyke, J.K., Thulin, C.D., Jean, F. and Thomas, G. (1997). Activation of the furin endoprotease is a multiple-step process: requirements for acidification and internal propeptide cleavage. *EMBO J.* **16**, 1508-1518.

Arima, K., Yonezawa, H., Uchikoba, T., Shimada, M., Kaneda, M. (2000). Cleavage specificity of cucumisin, a serine protease, with synthetic substrates. *Phytochemistry* **54**, 451-454.

Bacic, A., Harris, P.J. and Stone, B.A. (1988). Structure and function of plant cell walls. In *The biochemistry of plants: a comprehensive treatise* ,vol. **14** (ed. Preiss, J.). Academic Press Inc., pp.297-371.

Ballinger, M.D., Tom, J. and Wells, J.A. (1996). Furilisin: a subtilisin BPN' variant engineered for cleaving tribasic substrates. *Biochemistry* **35**, 13579-13585.

Bao, W., O'Malley, D.M., Whetten, R. and Sederoff, R.R. (1993). A laccase associated with lignification in loblolly pine xylem. *Science* **260**, 672-674.

Baron-Epel, O., Gharyal, P.K. and Schindler, M. (1988). Pectins as mediators of wall porosity in soybean cells. *Planta* **175**, 389-395.

Barr, P.J. (1991). Mammalian subtilisins: the long-sought dibasic processing endoproteases. *Cell* **66**, 1-3.

Barrett, A.J. (1994) Classification of peptidases. *Methods Enzymol.* **244**, 1-15.

Barrett, A.J. and McDonald, J.K. (1986). Nomenclature: protease, proteinase and peptidase. *Biochem. J.* **237**, 935.

Bauer, W.D. (1977). Plant cell walls. In *The molecular biology of plant cells* (ed. Smith, H.). Blackwell Scientific Publications, pp. 6-23.

Bauer, W.D., Talmadge, K.W., Keegstra, K. and Albersheim, P. (1973). The structure of plant cell walls. II. The hemicellulose of the walls of suspension-cultured sycamore cells. *Plant Physiol.* **51**, 174-187.

Bent, A.F., Kunkel, B., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J. and Staskawicz, B.J. (1994). *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant resistance genes. *Science* **265**, 1856-1860.

Bentley, A.K., Rees, D.J., Rizza, C. and Brownlee, G.G. (1986). Defective propeptide processing of blood clotting factor IX caused by mutation of arginine to glutamine at position -4. *Cell* **45**, 343-348.

Betzel, C., Klupsch, S., Papendorf, G., Hastrup, S., Branner, S. and Wilson, K.S. (1992) Crystal structure of the alkaline proteinase savinase from *Bacillus lentus* at 1.4 Å resolution. *J. Mol. Biol.* **223**, 427-445.

Bevan, M., Mayer, K., White, O., Eisen, J.A., Preuss, D., Bureau, T., Salzberg, S.L. and Mewes, H. (2001). Sequence and analysis of the *Arabidopsis* genome. Current Opinion in Plant Biology **4**(2), 105-110.

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for preparing recombinant plasmid DNA. *Nucleic Acid Research* **7**(6), 1513-1523.

Blakely, L.M., Blakely, R.M., Colowit, P.M. and Eliot, D.S. (1988). Experimental studies on lateral root formation in radish seedling roots. II. Analysis of the dose-response to exogenous auxin. *Plant Physiol.* **87**, 414-419.

Blaschek, W., Kohler, H., Senler, U. and Franz, G. (1982). Moleclar weight distribution of cellulose in primary cell walls – investigations with regenerating protoplasts, suspension cultured cells and mesophyll of tobacco. *Planta* **154**, 550-555.

Blow, D.M, Birktoft, J.J. and Hartley, B.S. (1969). Protein role of a buried acid group in the mechanism of action of chymotrypsin. *Nature* **221**, 337-340.

Bogacheva, A.M., Rudenskaya, G.N., Preusser, A., Tchikileva, I.O., Dunaevsky, Y.E., Golovkin, B.N. and Stepanov, V.M. (1999). A new subtilisin-like proteinase from roots of the dandelion *Taraxacum officinale* Webb S.L. *Biochemistry (Moscow)* **64**, 1030-1037.

Bogacheva, A.M. (1999). Plant subtilisins. *Biochemistry (Moscow)* **64**, 287-293.

Bolwell, G.P. (1993). Dynamic aspects of the plant extracellular matrix. *Int. Rev. Cytol.* **146**, 261-324.

Bond, J.S. (1996). Commercially available proteases. In *Proteolytic Enzymes - a practical approach* (eds. Benyon, R.J. and Bond, J.S.). Oxford University Press, pp.240.

Bowles, D.J. (1990). Defense-related proteins in higher plants. *Annu. Rev. Biochem.* **159**, 873-907

Bradford, M.M. (1976). A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.

Bradley, D.J., Kjellbom, P. and Lamb, C.J. (1992). Elicitor- and wound-induced oxidative cross-linking of a proline-rich plant cell wall protein: a novel, rapid defence response. *Cell* **70**, 21-30.

Brannigan, J.A., Dodson, G., Duggleby, H.J., Moody, P.C., Smith, J.L., Tomchick, D.R. and Murzin A.G. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* **378**, 416-419.

Brett, C.T. and Waldron, K.W. (1996). *Physiology and biochemistry of plant cell walls.* Topics in plant functional biology: **1** ,(eds. Black, M. and Charlwood, B.). Chapman and Hall, London.

Bronner-Fraser, M. (1990). Cell interactions in development. In *Seminars in Developmental Biology*, vol 1, Saunders Scientific Publications, Philadelphia, pp. 3-45.

Brown, J.A. and Fry, S.C. (1993). Novel *O*-D-galacturonosyl esters in the pectic polysaccharides of suspension-cultured plant cells. *Plant Physiol.* **103**, 993-999.

Brown, R.M., Jr. and Montezinos, D. (1976). Cellulose microfibrils: visualization of biosynthetic and orienting complexes in association with the plasma membrane. *Proc. Natl. Acad. Sci. USA* **73**, 143-147.

Burke, D., Kaufman, P., McNeil, M. and Albersheim, P. (1974). The structure of plant cell walls VI. A survey of the walls of suspension-cultured monocots. *Plant Physiol.* **54**, 109-115.

Butler, P.E., McKay, M.J. and Bond, J.S. (1987). Characterization of meprin, a membrane-bound metalloendopeptidase from mouse kidney. *Biochem J* **241**, 229-235.

Carpita, N.C. (1983). Hemicellulosic polymers of cell walls of Zea coleoptiles. *Plant Physiol.* **72**, 515-521.

Carpita, N.C. (1992). A special importance of the primary cell wall in plant development. In *Plant physiology*, Wadsworth Publishing, Belmont, CA, pp.334-336.

Carpita, N.C. (1996). Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 445-476.

Carpita, N.C., Sabularse, D., Montezinos, D. and Delmer, D.P. (1979). Determination of the pore size of cell walls of living plant cells. *Science* **205**, 1144-1147.

Carpita, N.C. and Gibeaut, D.M. (1993). Structural models of primary cell walls in flowering plants: cosistency of molecular structure with the physical properties of the walls during growth. *Plant J.* **3**(1), 1-30.

Cassab, G.I. and Varner, J.E. (1988). Cell wall proteins. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **39**, 321-353.

Cervantes, E., Rodriguez, A. and Nicolas, G. (1994). Ethylene regulates the expression of a cysteine proteinase gene during germination of chickpea (*Cicer arientinum*). *Plant Molecular Biology* **25**, 207-215.

Chambat, G., Joselau, J.P. and Barnoud, F. (1981). The carbohydrate constituents of the cell wall of suspension cultures of *Rosa glauca. Phytochemistry* **20**, 241-244.

Chang, C., Schaller, G.E., Patterson, S.E., Kwok, S.F., Meyerowitz, E.M. and Bleecker, A.B. (1992). The *TMK1* gene from *Arabidopsis* codes for a protein with structural and biochemical characteristics of a receptor protein kinase. *Plant Cell* **4**, 1263-1271.

Chang, T.K., Jackson, D.Y., Burnier, J.B. and Wells, J.A. (1994). Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl Acad. Sci. USA* **91**, 12544-12548.

Chen, K. and Arnold, F.H. (1993). Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc. Natl Acad. Sci. USA* **90**, 5618-5622.

Chen, W.T. (1992). Membrane proteases: roles in tissue remodelling and tumour invasion. *Curr. Opin. Cell Biol.* **4**, 802-809.

Chretien, M. and Li, C.H. (1967). Isolation, purification, and characterization of gamma-lipotropic hormone from sheep pituitary glands. *Can. J. Biochem.* **45**:1163-74

Chua, G.K. and Bushuk, W. (1969). Purification of wheat proteases by affinity chromatography on hemoglobin-Sepharose column. *Biochem Biophys Res Commun* **37**, 545-550.

Churms, S.C., Merrifield, E.H., Stephen, A.M., Walwyn, D.R., Polson, A., Van Der Merwe, K.F., Spies, H.S.C. and Costa, N. (1983). An L-arabinan from apple juice concentrates. *Carbohydrate Res.* **113**, 339-344.

Clarke, A.E. and Stone, B.A. (1988). Chemistry and biology of 1,3-β-glucans. Academic Press, Sydney.

Cleland, R.E. (1981). Wall extensibility. In *Encyclopaedia of plant physiology,* vol. **13B**, (eds. Tanner, W. and Loewus, F.A.). Springer-Verlag, Berlin.

Cosgrove, D.J. (1993). Wall extensibility: its nature, measurement and relationship to plant cell growth. *New Phytol.* **124**, 1-23.

Cosgrove, D.J. (1998). Cell wall loosening by expansins. *Plant Physiol.* **118**, 333-339.

Cosgrove, D.J. (1999). Enzymes and other agents that enhance cell wall extensibility. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 391-417.

Creemers, J.W., Groot Kormelink, P.J., Roebroek, A.J., Nakayama, K. and Van de Ven, W.J. (1993). Proprotein processing activity and cleavage site selectivity of the Kex2-like endoprotease PACE4. *FEBS Letts.* **336**, 65-69.

Curotto, E., Gonzalez, G., O'Reilly, S. and Tapia, G. (1989). Isolation and partial characterization of a protease from *Cucurbita ficifolia. FEBS Lett.* **243**, 363-365.

Cyr, R.J. (1991). Microtubule-associated protein in higher plants. In *The cytoskeletal basis of plant growth and form* (ed. Lloyd, C.W.). Academic Press, London, pp. 57-67.

Darvill, A.G., Smith, C.J. and Hall, M.A. (1978). Cell wall structure and elongation growth in *Zea mays* coleoptile tissue. *New Phytol.* **80**, 503-516.

Delmer, D.P. (1987). Cellulose biosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **38**, 259-290.

Delmer, D.P. (1991). The biochemistry of cellulose synthesis. In *The cytoskeletal basis of plant growth and form* (ed. Lloyd, C.W.). Academic Press, London, pp.101-107.

Delmer, D.P. and Amor, Y. (1995). Cellulose biosynthesis. *Plant Cell* **7**, 987-1000.

Delmer, D.P. and Stone, B.A. (1988). Biosynthesis of plant cell walls. In *The biochemistry of plants: a comprehensive treatise*, vol.**14** (ed. Stumpf, P.K. and Conn, E.E.). Academic Press, Inc., pp.373-420.

DeMartino, G.N. and Croall, D.E. (1983). Purification and characterization of a calcium-dependent protease from rat liver. *Biochemistry* **22**, 6287-6291.

Dice, J.F. (1987). Molecular determinants of protein half-lives in eukaryotic cells. *FASEB J.***1**, 349-57.

Dinwoodie, J.M. (1975).Timber: a review of the structure-mechanical property relationship. *J. Microsc.* **104**, 3-32.

Dixon, R.A. and Lamb, C.J. (1990). Molecular communication in interactions between plants and microbial pathogens. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **41**, 339-367.

Driouich, A., Faye, L. and Staehelin, L.A. (1993).The plant Golgi apparatus: a factory for complex polysaccharides and glycoproteins. *TIBS* **18**, 210-214.

D'Silva, I., Poirier, G.G. and Heath, M.C. (1998). Activation of cysteine proteases in cowpea plants during the hypersensitive response – a form of programmed cell death. *Experimental Cell Research* **245**, 389-399.

Dunn, B.M. (1996). Determination of protease mechanism. In *Proteolytic Enzymes - a practical approach* (eds. Benyon, R.J. and Bond, J.S.). Oxford University Press, pp.57-81.

Edwards, M., Dea, I.C.M., Bulpin, P.V. and Reid, J.S.G. (1985). Xyloglucan (amyloid) mobilization in the cotyledons of *Tropaeolum majus* L. seeds following germination. *Planta* **163**, 133-140.

Ellis, R.J. (1998). Steric chaperones. *TIBS* **23**, 43-45.

Estell, D.A., Graycar, T.P. and Wells, J.A. (1985). Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. *J. Biol. Chem.* **260**, 6518-6521.

Everdeen, D.S., Kiefer, S., Willard, J.J., Muldoon, E.P., Dey, P.M., Li, X.-B. and Lamport, D.T.A. (1988). Enzymic cross-linkage of monomeric extensin precursors in vitro. *Plant Physiol.* **87**, 616-621.

Feinberg, A.P. and Vogelstein, B. (1984). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6.

Feldmann, K.A. (1991). T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *Plant J.* **1**, 71-82.

Fincher, G.B., Stone, B.A. and Clarke, A.E. (1983). Arabinogalactan proteins: structure, biosynthesis and function. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **34**, 47-70.

Fisher, M.L., Anderson, A.J. and Albersheim, P. (1973). Host-pathogen interactions. VI. A single plant protein efficiently inhibits endopolygalacturonases secreted by *Colletotrichum lindemuthianum* and *Aspergillus niger*. *Plant Physiol.* **51**, 489-491.

Fleming, A.J., Caderas, D., Wehrli, E., McQueen-Mason, S. and Kuhlemeier, C. (1999). Analysis of expansin-induced morphogenesis on the apical meristem of tomato. *Planta* **208**, 166-174.

Florack, D.E. and Stiekema, W.J. (1994). Thionins: properties, possible biological roles and mechanisms of action. *Plant Mol Biol* **26**, 25-37.

Freedman, R.B., Brockway, B.E. and Lambert, N. (1984). Protein disulphide-isomerase and the formation of native disulphide bonds. *Biochem. Soc. Trans.* **12**, 929-932.

Fry, S.C. (1979). Phenolic components of the primary cell wall and their possible role in the hormonal regulation of growth. *Planta* **146**, 343-351.

Fry, S.C. (1982). Isodityrosine, a new cross-linking amino acid from plant cell wall glycoprotein. *Biochem. J.* **204**, 449-455.

Fry, S.C. (1986). Cross-linking of matrix polymers in the growing cell walls of angiosperms. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **37**, 165-186.

Fry, S.C. (1995). Polysaccharide-modifying enzymes in the plant cell wall. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 497-520.

Fry, S.C., Smith, R.C., Renwick, D.J., Hodge, S.K. and Matthews, K.J. (1992). Xyloglucan endotransglycosylase, a new wall-loosening enzyme activity from plants. *Biochem. J.* **282**, 821-828.

Fry, S.C., Aldington, S., Hetherington, P.R. and Aitken, J. (1993). Oligosaccharides as signals and substrates in the plant cell wall. *Plant Physiol.* **103**, 1-5.

Fry, S.C. (1998). Oxidative scission of plant cell wall polysaccharides by ascorbate-induced hydroxyl radicals. *Biochem. J.* **332**, 507-515.

Fuller, R.S., Sterne, R.E. & Thorner, J. (1988) Enzymes required for yeast prohormone processing. *Annu. Rev. Physiol.* **50**, 345-362.

Fuller, R.S., Brake, A. & Thorner, J. (1989) Yeast prohormone processing enzyme (KEX2 gene product) is a Ca2+ -dependent serine protease. *Proc. Natl Acad. Sci. USA* **86**, 1434-1438.

Giddings, T.H., Jr., Brower, D.L. and Staehelin, L.A. (1980). Visualization of particle complexes in the plasma membrane of *Micrasterias denticulata* associated with the formation of cellulose fibrils in primary and secondary cell walls. *J. Cell Biol.* **84**, 327-339.

Glotzer, M., Murray, A.W., and Kirschner, M.W. (1991). Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132-138.

Goldberg, R.B., Beals, T.P. and Sanders, P.M. (1993). Anther development: basic principles and practical applications. *Plant Cell* **5**, 1217-1229.

Granell, A., Belles, J.M. and Conejero, V. (1987). Induction of pathogenesis proteins in tomato by citrus exocortis viroid, silver ions and ethephon. *Mol. Plant Pathol.* **31**, 83-89.

Grunstein, M. and Hogness, D.S. (1975). Colony hybridization: a method for the isolation of cloned DNAs which contain a specific gene. *Proc. Natl. Acad. Sci. USA* **72**, 3961.

Guo, Z.J., Lamb, C. and Dixon, R.A. (1998). A serine protease from suspension-cultured soybean cells. *Phytochemistry* **47**, 547-53.

Halban, P.A. and Irminger, J. (1994). Sorting and processing of secretory proteins. *Biochem. J.* **299**, 1-18.

Hall, J.L. and Hawes, C. (1991). Electron microscopy of plant cells. Academic Press Ltd., London, pp. 186-192.

Hanks, S.K., Quinn, A.M. and Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**, 42-52.

Hardie, D.G. (1999). Plant protein serine/threonine kinases: classification and functions. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 1999 **50**, 97-131.

Hart, D.A. and Kindel, P.K. (1970). Isolation and partial characterization of apiogalacturonans from the cell wall of *Lemna minor*. *Biochem. J.* **116**, 569-579.

Hatsuzawa, K., Hosaka, M., Nakagawa, T., Nagase, M., Shoda, A., Murakami, K. and Nakayama, K. (1990). Structure and expression of mouse furin, a yeast Kex2-related protease. Lack of processing of coexpressed prorenin in GH4C1 cells. *J. Biol. Chem.* **265**, 22075-22078.

Hayashi, T. (1989). Xyloglucans in the primary cell wall. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 139-168.

Hayashi, T. (1991). Biochemistry of xyloglucans in regulating cell elongation and expansion. In *The cytoskeletal basis of plant growth and form* (ed. Lloyd, C.W.). Academic Press, London, pp. 131-144.

Hayashi, T., Kato, Y. and Matsuda, K. (1980). Xyloglucan from suspension-cultured soybean cells. *Plant Cell. Physiol.* **21**, 1405-1418.

Hayashi, T., Marsden, M.P.F. and Delmer, D. (1987). Pea xyloglucan and cellulose V. Xyloglucan-cellulose interactions *in vitro* and *in vivo*. *Plant Physiol.* **83**, 384-389.

Hetherington, P.R. and Fry, S.C. (1993). Xyloglucan endotransglycosylase activity in carrot cell suspensions during cell elongation and somatic embryogenesis. *Plant Physiol.* **103**, 987-992.

Heukeshoven, J. and Dernick, P. (1991). Electrophoresis Forum **186**, 22-27.

Higuchi, T. (1981). Biosynthesis of lignin. In *Encyclopaedia of plant physiology,* vol. **13B**, (eds. Tanner, W. and Loewus, F.A.). Springer-Verlag, Berlin, pp.194-224.

Himmelhoch, S.R. (1971). Chromatography of proteins on ion exchange adsorbents. *Meth. Enzymol.* **22**, 273-286.

Hiraiwa, N., Kondo, M., Nishimura, M. and Hara-Nishimura, I. (1997). An aspartic endopeptidase is involved in the breakdown of propeptides of storage proteins in protein-storage vacuoles of plants. *European Journal of Biochemistry* **246**, 133-141.

Hjerten, S., Rosengarten, J. and Pahlman, S. (1974). Hydrophobic interaction chromatography. The synthesis and use of some alkyl and aryl derivatives of agarose. *J Chromat.* **101**, 281-288.

Ho, I.K. and Hoskins, B. (1983). Variation of commercial diisopropyl fluorophosphate preparations in toxicological studies. *Drug Chem. Toxicol.* **6**(5), 421-427.

Horimoto, T., Nakayama, K., Smeekens, S.P. and Kawaoka, Y. (1994). Proprotein-processing endoproteases PC6 and furin both activate hemagglutinin of virulent avian influenza viruses. *J. Virol.* **68**(9), 6074-8.

Hoseney, R.C. and Faubion, J.M. (1981). *Cereal Chem.* **58**, 421-424.

Ikemura, H., Takagi, H. and Inouye, M. (1987). Requirement of pro-sequence for the production of active subtilisin E in *Escherichia coli. J. Biol. Chem.* **262**, 7859-7864.

Imagawa, H., Fukazawa, K. and Ishida, S. (1976). Study on the lignification in tracheids of Japanese larch, *Larix leptolepis* Gord. *Res. Bull. Col. Exp. Forests, Hokkaido Univ.* **33**, 1, 127-138.

Jackson, D.Y., Burnier, J., Quan, C., Stanley, M., Tom, J. and Wells, J.A. (1994) A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* **266**, 243-247.

James, G.T. (1978). Inactivation of the protease inhibitor phenylmethylsulphonyl fluoride in buffers. *Anal. Biochem.* **86**, 574-579.

Janzik, I., Macheroux, P., Amrhein, N. and Schaller, A. (2000). *Le*SBT1, a subtilase from tomato plants – overexpression in insect cells, purification and characterization. *J. Biol. Chem.* **275**, 5193-5199.

Jiang, L. and Rogers, J.C. (1999). Functional analysis of a Golgi-localized Kex2p-like protease in tobacco suspension culture cells. *Plant J.* **18**(1), 23-32.

John, P. (1992). Cellulose. In *Biosynthesis of the major crop products.* John Wiley and sons, pp.70-87.

Jolliffe, L.K., Doyle, R.J. and Streips, U.N. (1980). Extracellular proteases modify cell wall turnover in *Bacillus subtilis. J. Bacteriol.* **141**, 1199-1208.

Jordá, L., Coego, A., Conejero, V. and Vera, P. (1999). A genomic cluster containing four differentially regulated subtilisin-like processing protease genes is in tomato plants. *J. Biol. Chem.* **274**, 2360-2365.

Jordá, L., Conejero, V. and Vera, P. (2000). Characterization of *P69E* and *P69F*, two differentially regulated genes encoding new members of the subtilisin-like proteinase family from tomato plants. *Plant Physiol.* **122**, 67-73.

Jose, M. and Puigdomenech, P. (1993). Structure and expression of genes coding for structural proteins of the plant cell wall. *New Phytol.* **125**, 259-282.

Julius, D., Brake, A., Blair, L., Kunisawa, R. and Thorner, J. (1984). Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-α-factor. *Cell* **37**, 1075-1089.

Junghans, H. and Metzlaff, M. (1990). A Simple and Rapid Method for the Preparation of Total Plant DNA. *Biotechniques* **8**, 2, 176.

Juniper, B.E., Lawton, J.R. and Harris, P.J. (1981). Cellular organelles and cell wall formation in fibers from the flowering stem of *Lolium temulentum*. *New Phytol.* **89**, 609-619.

Kanaoka, Y., Takahashi, T., Nakayama, H. and Tanizawa, K. (1985). New fluorogenic substrates for subtilisin. *Chem Pharm Bull (Tokyo)* **33**, 1721-1724.

Kaneda , M. and Tominaga, N. (1975). Isolation and characterization of a proteinase from the sarcocarp of melon fruit. *J. Biochem.* **78**, 1287-1296.

Kaneda , M. and Tominaga, N. (1977). Isolation and characterization of proteinase from *Benincasa cerifera*. *Phytochemistry*. **16**, 345-346.

Kaneda, M. and Uchikoba, T. (1994). Protease from the sarcocarp of *Trichosanthes bracteata*. *Phytochemistry* **35**, 583-586.

Kaneda, M., Ohmine, H., Yonezawa, H. and Tominago, N. (1984). Amino acid sequence around the reactive serine of cucumisin from melon fruit. *J. Biochem.* **95**, 825-829.

Kaneda , M., Sobue, A., Eida, S. and Tominaga, N. (1986). Isolation and characterization of proteinases from the sarcocarp of snake-gourd fruit. *J. Biochem.* **99**, 569-577.

Keegstra, K., Talmadge, K.W., Bauer, W.D. and Albersheim, P. (1973). The structure of plant cell walls. III. A model of the walls of suspension-cultured sycamore cells based on interconnections of the macromolecular components. *Plant Physiol.* **51**, 188-196.

Keith, F.J. and Gay, N.J. (1990). The *Drosophila* membrane receptor Toll can function to promote cellular adhesion. *EMBO J.* **9**, 4299-4306.

Kido, H., Yokogoshi, Y., Sakai, K., Tashiro, M., Kishino, Y., Fukutomi, A. and Katunuma, N. (1992). Isolation and characterization of a novel trypsin-like protease found in rat bronchiolar epithelial Clara cells. A possible activator of the viral fusion glycoprotein. *J Biol Chem.* **267**, 13573-9.

Kieliszewski, M.J. and Lamport, D.T.A. (1994). Extensin: repetitive motifs, functional sites, post-translational codes and phylogeny. *Plant J.* **5**, 157-172.

Kobayashi, T., Kobayashi, E., Sato, S., Hotta, Y., Miyajima, N., Tanaka, A. and Tabata, S. (1994). Characterization of cDNAs induced in meiotic prophase in lily microporocytes. *DNA Res.* **1**, 15-26.

Kobe, B. and Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding domain. *TIBS* **19**, 415-421.

Kolattakuddy, P.E., Espelie, K.E. and Soliday, C.L. (1981). Hydrophobic layers attached to cell walls. Cutin, suberin and associated waxes. In *Encyclopaedia of plant physiology,* vol. **13B**, (eds. Tanner, W. and Loewus, F.A.). Springer-Verlag, Berlin, pp.225-254.

Kolattakuddy, P.E. (1980). Biopolyester membranes of plants: cutin and suberin. *Annu. Rev. Plant Physiol.* **32**, 539-567.

Konno, H., Yamaske, Y. and Katoh, K. (1986). Enzymatic degradation of pectic substances and cell walls purified from carrot cell cultures. *Phytochemistry* **25**, 623-627.

Kooiman, P. (1957). Amyloids of plant seeds. *Nature* **179**, 107-109.

Kristen, U. (1985). General and molecular cytology: the cell wall. *Progress in Botany* **103**, 1-5.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lamb, C.J. and Dixon, R.A. (1994). Molecular mechanisms underlying induction of plant defence gene transcription. *Biochem Soc Symp* **60**, 241-248.

Lamport, D.T.A. (1963). Oxygen fixation into hydroxyproline of plant cell wall protein. *J. Biol. Chem.* **238**, 1438-1440.

Lamport, D.T.A. and Epstein, L. (1983). A new model for the primary cell wall: a concatenated extensin-cellulose network. *Curr. Top. Plant Biochem. Physiol.* **2**, 73-83.

Landschulz, W.H., Johnson, P.F, and McKnight, S.L. (1988). The leucine zipper: a hypothetical structure common to new DNA binding proteins. *Science* **240**, 1795-1764.

Lau, J.M., McNeil, M., Darvill, A.G. and Albersheim, P. (1985). Structure of the backbone of rhamnogalacturonan I, a pectic polysaccharide in the primary walls of plants. *Carbohydrate Res.* **137**, 111-125.

Lawton, M. A. and Lamb, C.J. (1987). Transcriptional activation of plant defence genes by fungal elicitor, wounding, and infection. *Mol Cell Biol* **7**, 335-341.

Leach, J.E., Cantrell, M.A. and Sequeira, L. (1982). Hydroxyproline-rich bacterial agglutinin from potato: extraction, purification and characterization. *Plant Physiol.* **70**, 1355-1358.

Leibowitz, M.J. and Wickner, R.B. (1976) A chromosomal gene required for killer plasmid expression, mating, and spore maturation in *Saccharomyces cerevisiae. Proc. Natl Acad. Sci. USA* **73**, 2061-2065.

Leighton, T.J., Dor, R.H., Warren, R.A. and Kelln, R.A. (1973). The relationship of serine protease activity to RNA polymerase modification and sporulation in *Bacillus subtilis. J. Mol. Biol.* **76**, 103-122.

Levy, S., York, W.S. and Stuike-Prill, R. (1991). Simulations of static and dynamic molecular configurations of xyloglucan- the role of the fucosylated side chain in sugar-specific side chain folding. *Plant J.* **1**, 195-215.

Lewis, N.G. and Yamamoto, E. (1990). Lignin: occurrence, biogenesis and biodegradation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **41**, 455-496.

Linderstrøm-Lang, K. and Ottensen, M. (1947) A new protein from ovalbumin. *Nature* **159**, 807.

Loewus, F.A., Baldi, B.G., Franceschi, V.R., Meinert, L.D. and McCollum, J.J. (1985). Pollen sporoplasts – dissolution of pollen walls. *Plant Physiol.* **78**, 652-654.

Lopez, J.A., Chung, P.W., Fujikawa, K., Hagen, F.S., Davie, E.W. and Roth, G.J. (1987). Cloning of the α-chain of human platelet glycoprotein 1b: a transmembrane protein with homology to leucine-rich α-glycoprotein. *Proc. Natl. Acad. Sci. USA* **84**, 5615-5619.

Lowry, O.J., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951). Protein measurement with the folin-phenol reagent. *J.Biol. Chem.* **193**, 265-275.

Lusson, J., Vieau, D., Hamelin, J., Day, R., Chrétien, M. and Seidah, N.G. (1993). cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a candidate proprotein convertase expressed in endocrine and non-endocrine cells. *Proc. Natl Acad. Sci. USA* **90**, 6691-6695

Lusson, J., Benjannet, S., Hamelin, J., Savaria, D., Chretien, M. and Seidah, N.G. (1997). The integrity of the RRGDL sequence of the proprotein convertase PC1 is critical for its zymogen and C-terminal processing and for its cellular trafficking. *Biochem. J.* **326**, 737–744

McCann, M.C. Wells, B. and Roberts, K. (1990). Direct visualization of cross-links in the primary cell wall. *J. Cell Sci.* **96**, 323-334.

McCann, M.C. and Roberts, K. (1991). Architecture of the primary cell wall. In *The cytoskeletal basis of plant growth and form* (ed. Lloyd, C.W.). Academic Press, London, pp. 109-129.

McNeil, M., Albersheim, P., Taiz, L. and Jones, R.L. (1975). The structure of plant cell wall.

Plant Physiol. **55**, 64-68.

McNeil, M., Darvill, A.G., Fry, S.C. and Albersheim, P. (1984). Structure and function of the primary cell walls of plants. *Annu. Rev. Biochem.* **53**, 625-663.

McQueen-Mason, S.J., Durachko, D.M. and Cosgrove, D.J. (1992). Two endogenous proteins that induce cell wall extension in plants. *Plant Cell* **4**, 1425-1433.

McQueen-Mason, S.J., Fry, S.C., Durachko, D.M. and Cosgrove, D.J. (1993). The relationship between xyloglucan endotransglycosylase and *in vitro* cell wall extension in cucumber hypocotyls. *Planta* **190**, 327-331.

McQueen-Mason, S.J. and Cosgrove, D.J. (1994). Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *Proc. Natl. Acad. Sci. USA* **91**, 6574-6578.

McQueen-Mason, S.J. and Cosgrove, D.J. (1995). Expansin mode of action on cell walls: analysis of wall hydrolysis, stress relaxation and binding. *Plant Physiol.* **107**, 87-100.

Mains, R.E., Berard, C.A., Denault, J.B., Zhou, A., Johnson, R.C. and Leduc, R. (1997). PACE4: a subtilisin-like endoprotease with unique properties. *Biochem. J.* **321**, 587-593.

Mankarios, A.J., Hall, M.A., Jarvis, M.C., Threlfall, D.R. and Friend, J. (1980). Cell wall polysaccharides from onions. *Phytochemistry* **19**, 1731-1733.

Markland, F.S. and Smith, E.L. (1971). Subtilisins: primary structure, chemical and physical properties. In: *The Enzymes*, vol. 3 (ed. Boyer, P.D.). London: Academic Press, pp. 561-608.

Marx-Figini, M. and Schulz, G.V. (1966). On kinetics and mechanism of biosynthesis of cellulose in higher plants (according to experiments on cottonseed fibers). *Biochim. Biophys. Acta* **112**, 81-101.

Mattysse, A.G., White, A.R. and Lightfoot, R. (1995). Genes required for cellulose synthesis in *Agrobacterium tumefaciens*. *J. Bacteriol.* **177**, 1069-1075.

Mattysse, A.G., Thomas, D.L. and White, A.R. (1995). Mechanism of cellulose synthesis in *Agrobacterium tumefaciens*. *J. Bacteriol.* **177**, 1076-1081.

Maurizi, M.R. (1992). Proteases and protein degradation in *Escherichia coli*. *Experientia* **48**, 178-201.

Mayer, F. (1984). Special cytology: cytology and morphogenesis of the prokaryotic cell. *Progress in botany* **46**, 1-12.

Mbikay, M., Tadros, H., Lerner, C.P., Chen, A., El-Alfy, M., Clermont, Y., Seidah, N.G., Chrétien, M. and Simpson, E.M. (1997) Impaired fertility in mice deficient for the testicular germ-cell protease PC4. *Proc. Natl Acad. Sci. USA* **94**, 6842-6846.

Meagher, R.B. and Williamson, R.E. (1994). The plant cytoskeleton. In *Arabidopsis*, (eds. Meyerowitz, E. and Somerville, C.R.). Cold Spring Harbor Laboratory Press, NY, pp.1049-1084.

Meichtry, J., Amrhein, N. and Schaller, A. (1999). Characterisation of the subtilase gene family in tomato (*Lycopersicon esculentum* Mill.). *Plant Mol. Biol.* **39**, 749-760.

Meinert, M.C. and Delmer, D.P. (1977). Changes in biochemical composition of the cell wall of the cotton fibre during development. *Plant Physiol.* **59**, 1088-1097.

Meinkoth, J. and Wahl, G. (1984). Hybridization of nucleic acids immobilized on solid supports. *Anal. Biochem.* **138**, 267-284.

Meyer, K., Kohler, A. and Kauss, H. (1991). Biosynthesis of ferulic acid esters of plant cell wall polysaccharides in endomembranes from parsley cells. *FEBS Lett.* **290**, 209-212.

Meyerowitz, E.M. (1989). *Arabidopsis*, a useful weed. *Cell* **56**, 263-269.

Meyerowitz, E.M. and Pruitt, R.E. (1985). *Arabidopsis thaliana* and plant molecular genetics. *Science* **229**, 1214-1218.

Mizuno, K., Nakamura, T., Takada, K., Sakakibara, S. and Matsuo, H. (1987). A membrane-bound, calcium-dependent protease in yeast alpha-cell cleaving on the carboxyl side of paired basic residues. *Biochem Biophys Res Commun* **144**, 807-814.

Mizuno, K., Nakamura, T., Ohshima, T., Tanaka, S. and Matsuo, H. (1988). Yeast KEX2 gene encodes an endopeptidase homologous to subtilisin-like serine proteases. *Biochem. Biophys. Res. Comm.* **156**, 246-254.

Molloy, S.S., Thomas, L., Van Slyke, J.K., Stenberg, P.E. and Thomas, G. (1994). Intracellular trafficking and activation of the furin proprotein convertase: localization to the TGN and recycling from the cell surface. *EMBO J.* **13**, 18-33.

Moore, P.J., Darvill, A.G., Albersheim, P. and Staehelin, L.A. (1986). Immunogold localization of xyloglucan and rhamnogalacturonan I in the cell walls of suspension-cultured sycamore cells. *Plant Physiol.* **82**, 787-794.

Morihara, K. and Tsuzuki, H. (1969). Comparison of the specificities of various serine proteinases from microorganisms. *Arch. Biochem. Biophys.* **129**, 620-634.

Morris, E.R., Powell, D.A., Gidley, M.J. and Rees, D.A.(1982). Confirmation and interactions of pectins. *J.Mol.Biol.* **155**, 507-516.

Mort and Lamport (1977). Anhydrous hydrogen fluoride deglycosylates glycoproteins. *Anal. Biochem.* **82**, 289-309.

Mueller, S.C. and Brown, R.M. Jr. (1980). Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants. *J. Cell Biol.* **84**, 315-326.

Munzer, J.S., Basak, A., Zhong, M., Mamarbachi, A., Hamelin, J., Savaria, D., Lazure, C., Benjannet, S., Chretien, M. and Seidah, N.G. (1997). *In vitro* characterization of the novel proprotein convertase PC7. *J. Biol. Chem.* **272**, 19672-19681.

Mutlu, A., Chen, X., Reddy, S.M. and Gal, S. (1999). The aspartic proteinase expressed in *Arabidopsis thaliana* seeds and localized in the protein bodies. *Seed Science Research* **9**, 75-84.

Nakayama, K. (1997). Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing a wide variety of precursor proteins. *Biochem. J.* **327**, 625-635.

Nakayama, K., Kim, W.S., Torii, S., Hosaka, M., Nakagawa, T., Ikemizu, J., Baba, T. and Murakami, K. (1992). Identification of the fourth member of the mammalian endoprotease family homologous to the yeast Kex2 protease - its testis-specific expression. *J. Biol. Chem.* **267**, 5897-5900.

Neurath, H. (1986). The versatility of proteolytic enzymes. *J. Cell Biochem.* **32**, 35-49.

Neuteboom, L.W., Ng, J.M.Y., Kuyper, M., Clijdesdale, O.R., Hooykaas, P.J.J. and van der Zaal, B.J. (1999). Isolation and characterisation of cDNA clones corresponding with mRNAs that accumulate during auxin-induced lateral root formation. *Plant Mol. Biol.* **39**, 273-287.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C. (1994). Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol.* **106**, 1241-1255.

Ohta, Y., Hojo, H., Aimoto, S., Kobayashi, T., Zhu, X., Jordan, F. and Inouye, M. (1991). Pro-peptide as an intramolecular chaperone: renaturation of denatured subtilisin E with a synthetic pro-peptide. *Mol Microbiol* **5**, 1507-1510.

O'Neill, M. and Selvendran, R.R. (1980). Glycoproteins from the cell wall of *Phaseolus coccineus*. *Biochem. J.* **187**, 53-63.

O'Neill, M., Albersheim, P. and Darvill, A. (1990). The pectic polysaccharides of primary cell walls. In *Methods in plant biochemistry* vol.**2**, (ed. Dey, P.M.). Academic Press Ltd., pp.415-441.

Ottensen, M. and Svensen, I. (1970). The subtilisins. *Methods Enzymol.* **19**, 199-215.

Pang, P.P. and Meyerowitz, E.M. (1987). *Arabidopsis thaliana*: a model system for plant molecular biology. *Biotechnology* **5**, 1177-1181.

Palevitz, B.A. (1986). Division plane determination in guard mother cells of *Allium*: video time-lapse analysis of nuclear movements and phragmoplast rotation in the cortex. *Dev. Biol.* **117**, 644-654.

Pantoliano, M.W., Ladner, R.C., Bryan, P.N., Rollence, M.L., Wood, J.F. and Poulos, T.L. (1987). Protein engineering of subtilisin BPN': enhanced stabilization through the introduction of two cysteines to form a disulfide bond. *Biochemistry* **26**, 2077-2082.

Pantoliano, M.W., Whitlow, M., Wood, J.F., Rollence, M.L., Finzel, B.C., Gilliland, G.L., Poulos, T.L. and Bryan, P.N. (1988). The engineering of binding affinity at metal binding sites for the stabilization of proteins: subtilisin as a test case. *Biochemistry* **27**, 8311-8317.

Pantoliano, M.W., Whitlow, M., Wood, J.F., Dodd, S.W., Hardman, K.D., Rollence, M.L. and Bryan, P.N. (1989). Large increases in general stability for subtilisin BPN¢ through incremental changes in the free energy of unfolding. *Biochemistry* **28**, 7205-7213.

Pearce, G., Johnson, S. and Ryan, C.A. (1993). Structure-activity of deleted and substituted systemin, an 18-amino acid polypeptide inducer of plant defensive genes. *J. Biol. Chem.* **268**, 212-216.

Pennell R.I. and Lamb C. (1997). Programmed cell death in plants. *Plant Cell* **9**, 1157-1168.

Percival, E. and McDowell, R.H. (1981). Algal walls- composition and biosynthesis. In *Encyclopaedia of plant physiology,* vol. **13B**, (eds. Tanner, W. and Loewus, F.A.). Springer-Verlag, Berlin.

Perona, J.J. and Craik, C.S. (1995) Structural basis of substrate specificity in the serine proteases. *Protein Sci.* **4** , 337-360.

Piggot, P.J. and Coote, J.G. (1976). Genetic aspects of bacterial endospore formation. *Bacteriol. Rev.* **40**, 908-962.

Polgár, L. (1987). Structure and function of serine proteases. In *Hydrolytic Enzymes*, (eds. Neuberger, A. and Brocklehurst, K.), Elsevier Science Publishers B.V., pp.159-199.

Potter, I. and Fry, S.C. (1993). Xyloglucan endotransglycosylase activity in pea internodes. *Plant Physiol.* **103**, 235-241.

Priest, F.G. (1977). Extracellular enzyme synthesis in the genus *Bacillus. Bacteriol. Rev.* **41**, 711-753.

Rawlings, N.D. (1997). Serine and threonine peptidases. In *Handbook of Proteolytic Enzymes*, (eds. Barrett, A.J., Rawlings, N.D. and Woessner, J.F.), Academic Press Ltd.

Rawlings, N.D. and Barrett, A.J. (1994). Families of serine peptidases. *Methods Enzymol.* **244**, 19-61.

Rayle, D.L. and Cleland, R.E. (1992). The acid growth theory of auxin-induced cell elongation is alive and well. *Plant Physiol.* **99**, 1271-1274.

Reiter, W.-D. (1994). Structure, synthesis and function of the plant cell wall. In *Arabidopsis*, (eds. Meyerowitz, E. and Somerville, C.R.). Cold Spring Harbor Laboratory Press, NY, pp.955-988.

Reiter, W.-D., Chapple, C.C.S. and Somerville, C.R. (1993). Altered growth and cell walls in a fucose-deficient mutant of *Arabidopsis*. *Science* **261**, 1032-1035.

Rheinnecker, M., Eder, J., Pandey, P.S. and Fersht, A.R. (1994) Variants of subtilisin BPN' with altered specificity profiles. *Biochemistry* **33**, 221-225.

Ribeiro, A., Akkermans, A.D.L., van Kammen, A., Bisseling, T. and Pawlowski, K. (1995). A nodule-specific gene encoding a subtilisin-like protease is expressed in early stages of actinorhizal nodule development. *Plant Cell* **7**, 785-794.

Rich, D.H., Moon, B.J. and Harbeson, S. (1984). Inhibition of aminopeptidases by amastatin and bestatin derivatives. Effect of inhibitor structure on slow-binding processes. *J. Med. Chem.* **27**(4), 417-422.

Riggs, C.D. and Horsch, A.P. (1995). Molecular cloning of an anther-specific gene from tomato. *Plant Physiol. (Suppl.)* **108**, 117.

Roberts, K. (1989). The plant extracellular matrix. *Current Opinions In Cell Biology* **1**, 1020-1027.

Robertson, D., Mitchell, G.P., Gilroy, J.S., Gerrish, C., Bolwell, G.P. and Slabas, A.R. (1997). Differential extraction and protein sequencing reveals major differences in patterns of primary cell wall proteins from plants. *J. Biol. Chem.* **272**, 15841-15848.

Ross, P., Mayer, R. and Benziman, M. (1991). Cellulose biosynthesis and function in bacteria. *Microbiol. Rev.* **55**, 35-58.

Rudenskaya, G.N., Stepanov, V.M., Zakharova, Yu.A., Revina, L.P. and Khodova, O.M. (1987). A thiol-dependent serine proteinase from sunflower leaves: a bacterial serine proteinase analogue. *Biokhimiya* **52**, 1753-1755.

Rudenskaya, G.N., Bogdanova, E.A., Revina, L.P., Golovkin, B.N. and Stepanov, V.M. (1995). Macluralisin- a serine proteinase from fruits of *Maclura pomifera* (Raf.) Schneid. *Planta* **196**, 174-179.

Rudenskaya, G.N., Bogacheva, A.M., Preusser, A., Golovkin, B.N. and Stepanov, V.M. (1998). Taraxalisin - subtilisin-like protease from dandelion – *Taraxacum officinale* Webb S.L. roots. *Faseb J.* **11**, 3147.

Russell, A.J. and Fersht, A.R. (1987) Rational modification of enzyme catalysis by engineering surface charge. *Nature* **328**, 496-500.

Ryan, C.A. (1978). Proteinase inhibitors in plant leaves: a biochemical model for pest-induced natural plant protection. *Trends Biochem. Sci.* **3**, 148-150.

Ryan, C.A. (1987). Oligosaccharide signalling in plants. *Annu. Rev. Cell Biol.* **3**, 295-317.

Ryan, C.A. and Farmer, E.E. (1991). Oligosaccharide signals in plants: a current assessment. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **42**, 651-674.

Sakai, J., Rawson, R.B., Espenshade, P.J., Cheng, D., Seegmiller, A.C., Goldstein, J.L. and Brown, M.S. (1998). Molecular identification of the sterol-regulated luminal protease that cleaves SREBPs and controls lipid composition of animal cells. *Mol Cell* **2**, 505-514.

Salisbury, F.B. and Ross, C.W. (1992). *Plant physiology*. Wadsworth Publishing, Belmont, CA.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Scalbert, A., Monties, B., Lallemand, J.Y., Guitlet, E. and Rolando, C. (1985). Ether linkage between phenolic acids and lignin fractions from wheat straw. *Phytochemistry* **24**, 1359-1362.

Schaller, A and Ryan, C.A. (1994). Identification of a 50kDa systemin-binding protein in tomato plasma membranes having Kex2p-like properties. *Proc. Natl. Acad. Sci. USA* **91**, 11802-11806.

Schneider, R. and Schweiger, M. (1991). A novel modular mosaic of cell adhesion motifs in the extracellular domains of the neurogenic *trk* and *trkB* tyrosine kinase receptors. *Oncogene* **6**, 1807-1811.

Seidah, N.G., Chretien, M. and Day, R. (1994) The family of subtilisin/kexin like pro-protein and pro-hormone convertases: divergent or shared functions. *Biochimie* **76**, 197-209.

Seidah, N.G., Hamelin, J., Mamarbachi, M., Dong, W., Tadros, H., Mbikay, M., Chrétien, M. and Day, R. (1996). cDNA structure, tissue distribution and chromosomal localization of rat PC7, a novel mammalian proprotein convertase closest to yeast kexin-like proteinases. *Proc. Natl Acad. Sci. USA* **93**, 3388-3393.

Seidah, N.G., Mowla, S.J., Hamelin, J., Mamarbachi, A.M., Benjannet, S., Toure, B.B., Basak, A., Munzer, J.S., Marcinkiewicz, J., Zhong, M., Barale, J.C., Lazure, C., Murphy, R.A., Chretien, M. and Marcinkiewicz, M. (1999a). Mammalian subtilisin/kexin isozyme SKI-1: A widely expressed proprotein convertase with a unique cleavage specificity and cellular localization. *Proc Natl Acad Sci U SA* **96**, 1321-1326.

Seidah, N.G., Benjannet, S., Hamelin, J., Mamarbachi, A.M., Basak, A., Marcinkiewicz, J., Mbikay, M., Chretien, M. and Marcinkiewicz, M. (1999b). The subtilisin/kexin family of precursor convertases. Emphasis on PC1, PC2/7B2, POMC and the novel enzyme SKI-1. *Ann N Y Acad Sci* **885**, 57-74.

Selvendran, R.R. and O'Neill, M.A. (1987). Isolation and analysis of cell walls from plant material. *Methods Biochem. Anal.* **32**, 25-153.

Selvendran, R.R. and Ryden, P. (1990). Isolation and analysis of plant cell walls. In *Methods in plant biochemistry* vol. **2**, (ed. Dey, P.M.). Academic Press Ltd., pp.549-579.

Shcherban, T.Y., Shi, J., Durachko, D.M., Guiltinan, M.J., McQueen-Mason, S.J., Shieh, M. and Cosgrove, D.J. (1995). Molecular cloning and sequence analysis of expansins - a highly

conserved, multigene family of proteins that mediate cell wall extension in plants. *Proc. Natl. Acad. Sci. USA* **92**, 9245-9249.

Schechter, I., Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **127**, 157-162.

Shedletzky, E., Shmuel, M., Trainin, T., Kalman, S. and Delmer, D. (1992). Cell wall structure in cells adapted to growth on the cellulose-synthesis inhibitor 2,6-dichlorobenzonitrile: a comparison between two dicots and a monocot. *Plant Physiol.* **100**, 120-130.

Shennan, K.I., Taylor, N.A., Jermany, J.L., Matthews, G. and Docherty, K. (1995). Differences in pH optima and calcium requirements for maturation of the prohormone convertases PC2 and PC3 indicates different intracellular locations for these events. *J Biol Chem* **270**, 1402-1407.

Shinde, U. and Inouye, M. (1993). Intramolecular chaperones and protein folding. *TIBS* **18**, 442-446.

Shotton, D. (1971). The molecular architecture of the serine proteinases. In *Proc Int Res Con Proteinase Inhibitors*, (eds. Fritz, H. and Tschesche, H.), Walter de Gruyter, Berlin, pp. 47-55.

Showalter, A.M. (1993). Structure and function of plant cell wall proteins. *Plant Cell* **5**, 9-23.

Sidhu, J.S., Hoseney, R.C., Faubion, J. and Nordin, P. (1980). Reaction of [^{14}C]-cysteine with wheat flour water solubles under ultraviolet light. *Cereal Chem.* **57**, 380-382.

Siezen, R.J., de Vos, W.M., Leunissen, A.M. and Dijkstra, B.W. (1991). Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteases. *Protein Eng.* **4**, 719-737.

Siezen, R.J. and Leunissen, J.A.M. (1997). Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Science* **6**, 501-523.

Simon, M.M., Simon, H.G., Fruth, U., Epplen, J., Muller-Hermelink. H.K. and Kramer, M.D. (1987). Cloned cytolytic T-effector cells and their malignant variants produce an extracellular matrix degrading trypsin-like serine proteinase. *Immunology* **60**, 219-230.

Smith, C.A., Toogood, H.S., Baker, H.M., Daniel, R.M. and Baker, E.N. (1999). Calcium-mediated thermostability in the subtilisin superfamily: the crystal structure of *Bacillus* Ak.1 protease at 1.8 angstrom resolution. *J. Mol. Biol.* **294**, 1027-1040.

Smith, R. and Turk, V. (1974). Cathepsin D: rapid isolation by affinity chromatography on haemoglobin-agarose resin. *Eur. J. Biochem.* **48**, 245-254.

Solomon, M., Belenghi, B., Delledonne, M., Menachem, E. and Levine, A. (1999). The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *Plant Cell* **11**(3), 431-44.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**(3), 503-517.

Spitzer, N.C. (1970). Low resistance connections between cells in the developing anther of the lily. *J.Cell Biol.* **45**, 565-575.

Stahl M.L. and Ferrari E. (1984). Replacement of the *Bacillus subtilis* subtilisin structural gene with an *in vitro*-derived deletion mutation. *J. Bacteriol.* **158**, 411-418.

Stambaugh R., Brackett B.G. and Mastroianni L. (1969). Inhibition of *in vitro* fertilization of rabbit ova by trypsin inhibitors. *Biol Reprod.* **1,** 223-227.

Steiner, D.F., Cunningham, D., Spiegelmann, L. and Aten, B. (1967). Insulin biosynthesis: evidence for a precursor. *Science* **157**, 697-699.

Stevens, B.J.H. and Selvendran, R.R. (1984). Pectic polysaccharides of cabbage (*Brassica oleracea*). *Phytochemistry* **23**, 107-115.

Stotz, H.U., Contos, J.J.A., Powell, A.L.T., Benett, A.B. and Labavitch, J.M. (1994). Structure and expression of an inhibitor of fungal polygalacturonases from tomato. *Plant Mol. Biol.* **24**, 607-617.

Stroud, R.M., Krieger, M., Koeppe, R.E., Kossiakoff, A.A. and Chambers, J.L. (1975). Structure-function relationships in the serine proteases. In *Proteases and biological control* (eds. Reich, E., Rifkin, D.B. and Shaw, E.). Cold Spring Harbor, N.Y., pp. 13-32.

Taiz, L. (1984). Plant cell expansion: regulation of cell wall mechanical properties. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **35**, 585-657.

Talbott, L.D. and Ray, P.M. (1992). Molecular size and separability of pea cell wall polysaccharides. Implications for model of cell wall structure. *Plant Physiol.* **98**, 357-368.

Talmadge, K.W., Keegstra, K., Bauer, W.D. and Albersheim, P. (1973). The structure of plant cell walls. I. The macromolecular components of the walls of suspension-cultured sycamore cells with a detailed analysis of the pectic polysaccharides. *Plant Physiol.* **51**, 158-173.

Tatusova, T.A. and Madden, T.L. (1999). BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett* **174**, 247-250.

Taylor, A.A., Horsch, A., Rzepczyk, A., Hasenkampf, C.A. and Riggs, C.D. (1997). Maturation and secretion of a serine proteinase is associated with events of late microsporogenesis. *Plant J.* **12**(6), 1261-1271.

Taylor, I.E.P. and Atkins, E.D.T. (1985). X-ray diffraction studies on the xyloglucan from tamarind (*Tamarindus indica*) seeds. *FEBS Lett.* **181**, 300-302.

Taylor, M.A.J., Baker, K.C., Briggs, G.S., Connerton, I.F., Cummings, N.J., Pratt, K.A., Revell, D.F., Freedman, R.B. and Goodenough, P.W. (1995). Recombinant pro-regions from papain and papaya proteinase IV are selective high affinity inhibitors of the mature papaya enzymes. *Protein Engineering* **8**, 59-62.

Terada, I. Kwon, S.-T., Miyata, Y., Matsuzawa, H. and Ohta, T. (1990). Unique precursor structure of an extracellular protease, aqualysin I, with NH2- and COOH-terminal pro-sequences and its processing in *Escherichia coli*. *J. Biol. Chem.* **265**, 6576-6581.

Thomas, J.R., McNeil, M. Darvill, A.G. and Albersheim, P. (1987). Structure of plant cell walls. 29. Isolation and characterization of wall polysaccharides from suspension-cultured Douglas-fir cells. *Plant Physiol.* **83**, 659-671.

Tornero, P., Mayda, E., Gomez, M.D., Canas, L., Conejero, V. and Vera, P. (1996a). Characterisation of LRP, a leucine-rich repeat (LRR) protein from tomato plants that is processed during pathogenesis. *Plant J.* **10**, 315-330.

Tornero, P., Conejero, V. and Vera, P. (1996b). Primary structure and expression of a pathogen-induced protease (PR-P69) in tomato plants: similarity of functional domains to subtilisin-like endoproteases. *Proc. Natl. Acad. Sci.* **93**, 6332-6337.

Tornero, P., Conejero, V. and Vera, P. (1997). Identification of a new pathogen-induced member of the subtilisin-like processing protease family from plants. *J. Biol. Chem.* **272**, 14412-14419.

Twining, S.S. (1984). Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes. *Anal. Biochem.* **143**, 30-34.

Uchikoba, T. and Kaneda, M. (1996). Milk-clotting activity of cucumisin, a plant serine protease from melon fruit. *Appl. Biochem. Biotech.* **56**, 325-330.

Uchikoba, T., Horita, H. and Kaneda, M. (1990). Protease from the sarcocarp of yellow snake-gourd fruit. *Phytochemistry* **29**, 1879-1881.

Uchikoba, T., Niidome, T., Sata, I. and Kaneda, M. (1993). Protease D from the sarcocarp of honeydew melon fruit. *Phytochemistry* **33**, 1005-1008.

Uchikoba, T., Yonezawa, H. and Kaneda, M. (1995). Cleavage specificity of cucumisin, a plant serine protease. *J. Biochem. (Tokyo)* **117**, 1126-1130.

Uchikoba, T., Yonezawa, H., Kaneda, M. (1998). Cucumisin like protease from the sarcocarp of *Benincasa hispida* var. Ryukyu. *Phytochemistry* **49**, 2215-2219.

Valent, B.S. and Albersheim, P. (1974). The structure of plant cell walls V. On the binding of xyloglucan to cellulose fibres. *Plant Physiol.* 54, 105-108.

Vance, C.P., Kirk, T.K. and Sherwood, R.T. (1980). Lignification as a mechanism of disease resistance. *Annu. Rev. Phytopathol.* **18**, 259-288.

Varner, J.E. and Lin, L.-S. (1989). Plant cell wall architecture. *Cell* **56**, 231-239.

Vierstra, R.D. (1993). Protein degradation in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44**, 385-410.

Vierstra, R.D. (1996). Proteolysis in plants: mechanisms and functions. *Plant Mol. Biol.* **32**, 275-302.

Vera, P. and Conejero, V. (1988). Pathogenesis-related proteins of tomato – P69 as an alkaline endoprotease. *Plant Physiol.* **87**, 58-63.

Vera, P., Lamb, C. and Doerner, P.W. (1994). Cell-cycle regulation of hydroxyproline-rich glycoprotein *HRGPnt3* gene expression during the initiation of lateral root meristems. *Plant J.* **6**, 717-727.

Von Heinje, G. (1986). A new method for predicting signal sequence cleavage site. *Nucleic Acids Res* **14**, 4683-4690.

Wada, S. and Ray, P.M. (1978). Matrix polysaccharides of oat coleoptile cell walls. *Phytochemistry* **17**, 923-931.

Walker, J.C. (1993). Receptor-like protein kinase genes of *Arabidopsis thaliana. Plant J.* **3**, 451-456.

Walton, J.D. (1994). Deconstructing the cell wall. *Plant Physiol.* **104**, 1113-1118.

Wang, S. and Carpenter, F.H. (1967). Kinetics of the tryptic hydrolysis of the oxidized B chain of bovine insulin. *Biochemistry* **6**(1), 215-224.

Wardrop, A.B. and Bland, D.E. (1959). The process of lignification in woody plants. In *Biochemistry of woods* (eds. Kratzl, K and Billek, G.). Pergamon Press, New York, pp.92-116.

Wells, J.A. and Powers, D.B. (1986). *In vivo* formation and stability of engineered disulfide bonds in subtilisin. *J. Biol. Chem.* **261**, 6564-6570.

Wells, J.A. and Estell, D.A. (1988). Subtilisin B an enzyme designed to be engineered. *Trends Biochem. Sci.* **13**, 291-297.

Wells, J.A., Powers, D.B., Bott, R.R., Graycar, T.P. and Estell, D.A. (1987). Designing substrate specificity by protein engineering of electrostatic interactions. *Proc. Natl Acad. Sci. USA* **84**, 1219-1223.

Wessels, J.G.H. and Sietsma, J.H. (1981). Fungal cell walls: a survey. In *Encyclopaedia of plant physiology,* vol. **13B**, (eds. Tanner, W. and Loewus, F.A.). Springer-Verlag, Berlin, pp. 352-394.

Whitmore, F.W. (1981). Lignin-protein complex in cells walls of *Pinus elliottii*: amino acid constituents. *Phytochemistry* **21**, 315-318.

Wilcox, C.A. and Fuller, R.S. (1991) Posttranslational processing of the prohormone-cleaving Kex2 protease in the *Saccharomyces cerevisiae* secretory pathway. *J. Cell Biol.* **115**, 297-307.

Wilder, B.M. and Albersheim, P. (1973). The structure of plant cell walls. IV. A structural comparison of the wall hemicellulose of cell suspension cultures of sycamore (*Acer pseudopalantus*) and of red kidney bean (*Phaseolus vulgaris*). *Plant Physiol.* **51**, 889-893.

Wilkinshaw, M.D. and Arnott, S. (1981). Confirmation and interactions of pectins II. Models for junction zones in pectinic acid and calcium pectate gels. *J.Mol.Biol.* **153**, 1075-1085.

Wong, H.C., Fear, A.L., Calhoon, R.D., Eichinger, G.H., Mayer, R., Amikam, D., Benziman, M., Gelfand, D.H., Meade, J.H., Emerick, A.W., Bruner, R., Ben-Bassat, A. and Tal, R. (1990). Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. *Proc. Natl. Acad. Sci. USA*, **87**, 8130-8134.

Wu, Z.-P. and Hilvert, D. (1989). Conversion of a protease into an acyl transferase: selenolsubtilisin. *J. Am. Chem. Soc.* **111**, 4513-4514.

Yamagata, H., Masuzawa, T., Ueno, S. and Iwasaki, T. (1989). Isolation and characterization of a possible native cucumisin from developing melon fruits and its limited autolysis to cucumisin. *Agric. Biol. Chem.* **53**, 1009-1017.

Yamagata, H., Masuzawa, T., Nagaoka, Y., Ohnishi, T. and Iwasaki, T. (1994). Cucumisin, a serine protease from melon fruits, shares structural homology with subtilisin and is generated from a large precursor. *J. Biol. Chem.* **269**, 32725-32731.

Yamagata, H., Uesugi, M., Saka, K., Iwasaki, T. and Aizono, Y. (2000). Molecular cloning and characterization of a cDNA and a gene for subtilisin-like proteases from rice (*Oryza sativa* L.) and *Arabidopsis thaliana. Biosci. Biotech. Biochem.* **64** (9), 1947-1957.

Yamazaki, N., Fry, S.C., Darvill, A.G. and Albersheim, P. (1983). Host-pathogen interactions. 24. Fragments isolated from suspension-cultured sycamore cell walls inhibit the ability of the cells to incorporate [leucine-^{14}C] into proteins. *Plant Physiol.* **72**, 864-869.

Yang, J. and Kramer, J.M. (1994). *In vitro* mutagenesis of *Caenorhabditis elegans* cuticle collagens identifies a potential subtilisin-like protease cleavage site and demonstrates that carboxyl domain disulphide bonding is required for normal function but not assembly. *Mol. Cell. Biol.* **14**, 2722-2730.

Yang, M.Y., Ferrari, E. and Henner, D.J. (1984). Cloning of the neutral protease gene of *Bacillus subtilis* and the use of the cloned gene to create an *in vitro*-derived deletion mutation. *J. Bacteriol.* **160**, 15-21.

Ye, Z.H. and Varner, J.E. (1991). Tissue-specific expression of cell wall proteins in developing soybean tissues. *Plant Cell* **3**, 23-27.

Yonezawa, H., Uchikoba, T. and Kaneda, M. (1996). Identification of the reactive histidine of cucumisin, a plant serine protease: modification with peptidyl chloromethyl ketone derivative of peptide substrate. *J. Biochem.* **118**, 917-920.

Yonezawa, H., Uchikoba, T. and Kaneda, M. (1997). Substrate specificity of honeydew melon protease D, a plant serine endopeptidase. *Biosci Biotechnol Biochem* **61**, 1277-1280.

York, W.S., Darvill, A.G. and Albersheim, P. (1984). Inhibition of 2,4-dichlorophenoxyacetic acid-stimulated elongation of pea stem segments by a xyloglucan oligosaccharide. *Plant Physiol.* **75**, 295-297.

York, W.S., Darvill, A.G., McNeil, M., Stevenson, T.T. and Albersheim, P. (1986). Isolation and characterization of plant cell walls and cell wall components. *Methods In Enzymology* **118**, 1-40.

York, W.S., Oates, J.E., van Halbeek, H., Darvill, A.G. and Albersheim, P. (1988). Location of the *O*-acetyl substituents on a nonasaccharide repeating unit of sycamore extracellular xyloglucan. *Carbohydr. Res.* **173**, 113-132.

Zhao, C., Johnson, B.J., Kositsup, B. and Beers, E.P. (2000). Exploiting secondary growth in *Arabidopsis*. Construction of xylem and bark cDNA libraries and cloning of three xylem endopeptidases. *Plant Physiology* **123**, 1185-1196.

Zhu, X., Ohta, Y., Jordan, F. and Inouye, M. (1989). Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intramolecular process. *Nature* **339**, 483-484.